PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, 5/10, C07K 14/47, 16/18, G01N 33/566, A61K 38/17, A01K 67/027

A2

(11) International Publication Number: **WO 99/27088**

(43) International Publication Date: 3 June 1999 (03.06.99)

(21) International Application Number: PCT/CA98/01058

(22) International Filing Date: 19 November 1998 (19.11.98)

(30) Priority Data:

60/066,140	19 November 1997 (19.11.97)	US
09/123,549	28 July 1998 (28.07.98)	US

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(71) Applicant *(for all designated States except US)*: MOUNT SINAI HOSPITAL [CA/CA]; 600 University Avenue, Toronto, Ontario M5G 1X5 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: DONOVIEL, Dorit [US/CA]; 191 Roe Avenue, Toronto, Ontario M5M 2J1 (CA). BERNSTEIN, Alan [CA/CA]; 69 Lyndhurst Avenue, Toronto, Ontario M5R 2Z8 (CA).

(74) Agents: TANDAN, Susan, I. et al.; Van Zant & Associates, Suite 1407, 77 Bloor Street West, Toronto, Ontario M5S 1M2 (CA).

(54) Title: NOVEL GENE AND PROTEIN EXPRESSED IN NEURAL AND PANCREATIC TISSUES

(57) Abstract

A nucleic acid molecule, designated *Sel–1L* (Sel–1 like; also previously known as Hip–1), which is widely expressed in adult mouse and human tissues with particularly high levels in the pancreas and in the floor plate of the neural tube. Sel–1L has also been localized within intracellular vesicles. The gene maps to 14q24.3–31 in close proximity to the insulin diabetes locus (IDDML). The nucleic acid molecules encodes a protein designated "SEL–1L Protein".

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Title: Novel Gene and Protein Expressed in Neural and Pancreatic Tissues

FIELD OF THE INVENTION

The invention relates to nucleic acid molecules, proteins encoded by such nucleic acid molecules; and use of the proteins and nucleic acid molecules.

BACKGROUND OF THE INVENTION

Alzheimer Disease (AD) manifests as the degeneration of neurons in specific areas of the brain and the subsequent loss of certain cognitive functions. The etiology of this degeneration has been the subject of much speculation. Within the last two years, progress has been made in the elucidation of this chain of events. Genetic linkage of the disease to certain loci (on chromosomes XXI, I, and XIV) has identified at least three genes with mutations having direct consequences on the development of AD. One gene is the Amyloid Precursor Protein (*APP*) gene, and the other two are the presenilin 1 and 2 genes (*PS1, PS2*, respectively) that appear to be related to each other (reviewed by Selkoe (Science 275:630-631, 1997). While many familial AD cases can be accounted for by dominant mutations in any one of these genes, there are many other familial and sporadic AD cases whose genetic causes have not yet been identified [Sherrington et al. (Nature 375:754-760, 1995) and Alzheimer's Disease Collaborative Group (Nature Genetics 11:219-222)]. This suggests the possible involvement of other unidentified genes.

The *APOE* gene has been found to have an impact on the severity and the age of onset of AD but does not directly cause AD. The product of the *APOE* gene may bind and facilitate cholesterol transport across membranes (reviewed in Hyman and Tanzi (N Engl J Med 333:1242-7, 1995)). Its role in AD remains elusive.

It is now generally accepted that the accumulation of a peptide called A-beta (Ab) which is proteolytically cleaved from the APP protein in plaques in the vicinity of neurons has a role in the death of the neurons. Mutations in the *APP, PS1*, or *PS2* genes directly cause the accumulation of a fibrillogenic form of Ab (reviewed by John Hardy (P.N.A.S. 94:2095-7, 1997)). However, it is yet unclear whether this event is sufficient to cause AD, or whether other mitigating events occur.

Much research has focused on understanding the role of the presenilin genes and whether mutations in the genes have effects in the body other than elevating Ab. It is known that the presenilins are transmembrane proteins, and are localized to intracellular vesicles, the endoplasmic reticulum and the golgi apparatus (Reviewed by Christian Haass (Neuron 18:687-690, 1997)); all cellular components involved in protein trafficking and processing. This is consistent with their proposed role in the processing and/or secretion of Ab peptide. However, there is little known about the function of the presenilin genes in the cell. A significant clue to their function has resulted from the identification of a homolog of the presenilins (*Sel-12*) in a genetic screen in the worm *C. elegans* for modifiers of mutations in the *Lin-12* (notch) gene (Levitan and Greenwald (Nature 377:351-354, 1995)). *Sel-12* appeared to be a positive regulator of *Lin-12*, and this appears to be conserved in the mouse since ablation of PS1 reduces expression of *NOTCH 1* and results in a phenotype reminiscient of *NOTCH 1*-null mice (Grant and Greenwald (Dev. 124:637-644, 1997)). Thus, work in lower organisms places the presenilins in a genetic pathway specifying neuronal fate and this appears to be conserved in mammals. Whether this role is related to the effect of the malfunctioning presenilins in AD is unknown at this time. However, it is likely that genes involved in the same genetic pathway as *NOTCH* and the presenilins might have a bearing on AD.

In a similar genetic screen to the one which identified presenilin as a positive regulator of *LIN12 (NOTCH)*, a second worm gene called *Sel-1* was identified as a negative regulator of Notch. A point mutation or truncation in the gene appeared to result in a *SEL-1* null function and the mutation effected lin12p hypomorphic mutations. *Sel-1p* acts cell autonomously, it contains a functional signal sequence but is not found in the plasma membrane, and it is localized to intracellular vesicles (Grant and Greenwald (Dev. 124:637-644, 1997)). *Sel-1p* shares similarity with one known protein, hrd3p, a yeast protein which regulates the degradation of 3-hydroxy-3-methyl-glutaryl coenzyme A (HMG-CoA) Reductase--an enzyme involved in the cholesterol biosynthesis pathway (Hampton et al. (Mol. Biol. Cell 7:2029-44, 1996)). Taken together, this information suggests that Sel-1 controls the degradation of Notch. Since the presenilins and sel-1 are both in the notch pathway a homolog of sel-1 is likely to play an important role in human disease.

SUMMARY OF THE INVENTION

The present inventors have identified a nucleic acid molecule, designated *Sel-1L* or *Sel-1l* (*Sel-1* like; also previously known as Hip-1), which is widely expressed in adult mouse and human tissues with particularly high levels in the pancreas and in the floor plate of the neural tube. *Sel-1L* has also been localized within intracellular vesicles. The gene maps to 14q24.3-31 in close proximity to the insulin diabetes locus. The nucleic acid molecule encodes a protein designated "SEL-1L Protein".

Sel-1L maps to the same chromosomal interval as *PS1*, making it a candidate for an Alzheimer's Disease associated gene. *Sel-1L* may also play a role in the etiology of other neurodegenerative diseases such as Parkinson's Disease. *Sel-1L* is expressed in the floor plate of the neural tube during specification of the dopaminergic neurons via contact-mediated induction by the floor plate. Dopaminergic neurons are absent in Parkinson's Disease patients. *Sel-1L* also maps very close to a microsatellite marker (D14s67) that has been linked to a locus affecting Insulin Dependent Diabetes Mellitus (IDDML). The expression of *Sel-1L* in neural and pancreatic tissues also suggests that the gene may have a role in pancreatic cancer. Individuals exhibiting AD pathology in post-mortem examinations have been found to have a prevalence of pancreatic cancer over any other type of cancer and over the incidence of this tumor type in non-AD individuals (Burke et al. Alz. Dis, Assoc. Disor 8:22-28, 1994). *Sel-1L* may also have a role in fat and cholesterol metabolism which is a contributing factor in AD, and conditions such as coronary heart disease.

Broadly stated the present invention relates to an isolated nucleic acid molecule which comprises:

(i) a nucleic acid sequence encoding a protein having substantial sequence identity preferably at least 70% sequence identity, with the amino acid sequence of SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10;

(ii) a nucleic acid sequence encoding a protein comprising the amino acid sequence shown in SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10;

(iii) nucleic acid sequences complementary to (i);

(iv) a degenerate form of a nucleic acid sequence of (i);

(v) a nucleic acid sequence capable of hybridizing under stringent conditions to a nucleic acid sequence in (i), (ii) or (iii);

(vi) a nucleic acid sequence encoding a truncation, an analog, an allelic or species variation of a protein comprising the amino acid sequence of SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10; or

(vii) a fragment, or allelic or species variation of (i), (ii) or (iii).

Preferably, the purified and isolated nucleic acid molecule comprises:

(i) a nucleic acid sequence comprising the sequence of SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ. ID. NO. 7, or SEQ. ID. NO. 9 wherein T can also be U;

(ii) nucleic acid sequences complementary to (i), preferably complementary to the full nucleic acid sequence of SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ. ID. NO. 7, or SEQ. ID. NO. 9;

(iii) a nucleic acid capable of hybridizing under stringent conditions to a nucleic acid of (i) or (ii) and preferably having at least 18 nucleotides; or

(iv) a nucleic acid molecule differing from any of the nucleic acids of (i) to (iii) in codon sequences due to the degeneracy of the genetic code.

The invention also contemplates a nucleic acid molecule comprising a sequence encoding a truncation of the SEL-1L Protein, an analog, or a homolog of the SEL-1L Protein or a truncation thereof. (SEL-1L Protein and truncations, analogs and homologs of the SEL-1L Protein are also collectively referred to herein as "SEL-1L Related Proteins").

The nucleic acid molecules of the invention may be inserted into an appropriate expression vector, i.e. a vector that contains the necessary elements for the transcription and translation of the inserted coding sequence. Accordingly, recombinant expression vectors adapted for transformation of a host cell may be constructed which comprise a nucleic acid molecule of the invention and one or more transcription and translation elements linked to the nucleic acid molecule.

The recombinant expression vector can be used to prepare transformed host cells expressing SEL-1L Related Proteins. Therefore, the invention further provides host cells containing a recombinant molecule of the invention. The invention also contemplates transgenic non-human mammals whose germ cells and somatic cells contain a recombinant molecule comprising a nucleic acid molecule of the invention, in particular one which encodes an analog of the SEL-1L Protein, or a truncation of the SEL-1L Protein.

The invention further provides a method for preparing SEL-1L Related Proteins utilizing the purified and isolated nucleic acid molecules of the invention. In an embodiment a method for preparing a SEL-1L Related Protein is provided comprising (a) transferring a recombinant expression vector of the invention into a host cell; (b) selecting transformed host cells from untransformed host cells; (c) culturing a selected transformed host cell under conditions which allow expression of the SEL-1L Related Protein; and (d) isolating the SEL-1L Related Protein.

The invention further broadly contemplates an isolated SEL-1L Protein comprising the amino acid sequence as shown in SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10.

The SEL-1L Related Proteins of the invention may be conjugated with other molecules, such as proteins, to prepare fusion proteins. This may be accomplished, for example, by the synthesis of N-terminal or C-terminal fusion proteins.

The invention further contemplates antibodies having specificity against an epitope of a SEL-1L Related Protein of the invention. Antibodies may be labeled with a detectable substance and used to detect proteins of the invention in tissues and cells.

The invention also permits the construction of nucleotide probes which are unique to the nucleic acid molecules of the invention and/or to proteins of the invention. Therefore, the invention also relates to a probe comprising a nucleic acid sequence of the invention, or a nucleic acid sequence encoding a protein of the invention, or a part thereof. The probe may be labeled, for example, with a detectable substance and it may be used to select from a mixture of nucleotide sequences a nucleic acid molecule of the invention including nucleic acid molecules coding for a protein which displays one or more of the properties of a protein of the invention.

The invention still further provides a method for identifying a substance which binds to a protein of the invention comprising reacting the protein with at least one substance which potentially can bind with the protein, under conditions which permit the formation of complexes between the substance and protein and assaying for complexes, for free substance, or for non-complexed protein. The invention also contemplates methods for identifying substances that bind to other intracellular proteins that interact with a SEL-1L Related Protein. Methods can also be utilized which identify compounds which bind to SEL-1L gene regulatory sequences (e.g. promoter sequences).

Still further the invention provides a method for evaluating a compound for its ability to modulate the biological activity of a SEL-1L Related Protein of the invention. For example a substance which inhibits or enhances the interaction of the protein and a substance which binds to the protein may be evaluated. In an embodiment, the method comprises providing a known concentration of a SEL-1L Related Protein, with a substance which binds to the protein and a test compound under conditions which permit the formation of complexes between the substance and protein, and removing and/or detecting complexes.

Compounds which modulate the biological activity of a protein of the invention may also be identified using the methods of the invention by comparing the pattern and level of expression of the protein of the invention in tissues and cells , in the presence, and in the absence of the compounds.

Sel-1L Related Proteins contain a PTB domain binding site and therefore substances (e.g. peptides) and methods are provided for modulating the interaction of a PTB domain containing protein and a Sel-1L PTB domain binding site.

Sel-1L Related Proteins also contain PDZ domains and therefore substances (e.g. peptides) and methods are provided for modulating the interaction of a Sel-1L PDZ domain and a protein containing a PDZ domain binding motif.

The substances and compounds identified using the methods of the invention, and peptides of the invention may be used to modulate the biological activity of a SEL-1L Related Protein of the invention, and they may be used in the treatment of conditions such as Alzheimer's Disease, IDDML, cancer (e.g. pancreatic cancer or insulinomas), neurodegenerative diseases such as Parkinson's Disease, stroke, vascular dementia, and conditions requiring modulation of fat or cholesterol metabolism such as coronary heart disease. Accordingly, the substances and compounds may be formulated into compositions for administration to individuals suffering from one of these conditions.

Therefore, the present invention also relates to a composition comprising one or more of a protein of the invention, a peptide of the invention, or a substance or compound identified using the methods of the

invention, and a pharmaceutically acceptable carrier, excipient or diluent. A method for treating or preventing Alzheimer's Disease, diabetes, cancer (e.g. pancreatic cancer), neurodegenerative diseases such as Parkinson's Disease, stroke, vascular dementia, and conditions requiring modulation of fat or cholesterol metabolism such as coronary heart disease is also provided comprising administering to a patient in need thereof, a SEL-1L Related Protein of the invention, or a composition of the invention.

Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples while indicating preferred embodiments of the invention are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

DESCRIPTION OF THE DRAWINGS

The invention will be better understood with reference to the drawings in which:

Figure 1 shows an alignment of amino acid sequences of the Sel-1L protein, sel-1, and EST's from the X-Ref database;

Figure 2A shows a sequence alignment of SEL-1L and Sel-1;

Figure 2B shows a comparison of a C-terminal conserved region among SEL-1L family members;

Figure 3A is a blot showing mSEL-1L transcripts in adult tissues;

Figure 3B is a blot showing hSEL-1L transcripts in adult tissues

Figure 4A is a dark-field image of a cross-section through an E11.5 mouse embryo showing moderate levels of mSEL-1L expression in the DRG's and low levels throughout the neural tube;

Figure 4B is a dark-field image of a cross-section through an E11.5 embryo showing high expression in the floor plate of the neural tube and lack of expression in the notochord;

Figure 4C is a bright-field image of the same section shown in Figure 4B;

Figure 4D is a bright field closeup of the same section shown in Figure 4B and Figure 4C, showing the silver grains condensation over the midline on the apical side of the floor plate;

Figure 4E is a dark-field of a saggital section through an E14.5 embryo showing moderate expression of mSEL-1L in the epithelial layer of the villi of the gut and high expression in the pancreas;

Figure 5A is a Western blot confirming the SEL-1L sera;

Figure 5B is a Western blot showing endogenous SEL-1L in rat islet cells but not in human pancreatic adenocarcinoma cell lines;

Figure 5C shows immunofluorescence staining with RinM cells showing the cytoplasmic location of the SEL-1L protein;

Figure 6A shows fluorescence staining illustrating hybridization with a genomic *SEL-1L* probe to human metaphase chromosomes;

Figure 6B shows Dapi-stained chromosomes demonstrating that the signal in Figure 6A corresponds to the long arm of chromosome 14;

Figure 6C is a giemsa-band diagram of human chromosome 14 showing the clustering of the fluorescence signal in Figure 6A to an interval between 14q24.3 and 31;

Figure 7 shows an alignment of the mouse and human sel-1l amino acid sequences;

Figure 8 are graphs showing the results of Bia-core and fluorescence polarization analyses of dNUMB-PTB domain binding to Sel-1L "NLGpY"-containing peptide.

DETAILED DESCRIPTION OF THE INVENTION

1. Nucleic Acid Molecules of the Invention

As hereinbefore mentioned, the invention provides an isolated nucleic acid molecule having a sequence encoding SEL-1L Protein. The term "isolated" refers to a nucleic acid substantially free of cellular material or culture medium when produced by recombinant DNA techniques, or chemical reactants, or other chemicals when chemically synthesized. An "isolated" nucleic acid may also be free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid molecule) from which the nucleic acid is derived. The term "nucleic acid" is intended to include DNA and RNA and can be either double stranded or single stranded. In an embodiment, the nucleic acid molecule encodes SEL-1L Protein comprising the amino acid sequence as shown in SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10, preferably a nucleic acid molecule comprising the nucleic acid sequence as shown in SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ. ID. NO. 7, or SEQ. ID. NO. 9.

The invention includes nucleic acid sequences complementary to a nucleic acid encoding SEL-1L Protein comprising the amino acid sequence as shown in SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10, preferably the nucleic acid sequences complementary to the full nucleic acid sequence shown in SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ. ID. NO. 7, or SEQ. ID. NO. 9.

The invention includes nucleic acid molecules having substantial sequence identity or homology to the nucleic acid sequence as shown in SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ. ID. NO. 7, or SEQ. ID. NO. 9 or encoding proteins having substantial identity or similarity to the amino acid sequence shown in SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10.. Preferably, the nucleic acids have substantial sequence identity for example at least 70% nucleic acid identity; more preferably 80% nucleic acid identity; and most preferably at least 89% to 95% sequence identity.

"Identity" as known in the art and used herein, is a relationship between two or more amino acid sequences or two or more nucleic acid sequences, as determined by comparing the sequences. It also refers to the degree of sequence relatedness between amino acid or nucleic acid sequences, as the case may be, as determined by the match between strings of such sequences. Identity and similarity are well known terms to skilled artisans and they can be calculated by conventional methods (for example see Computational Molecular Biology, Lesk, A.M. ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W. ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A.M. and Griffin, H.G. eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G. Acadmeic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J. eds. M. Stockton Press, New York, 1991, Carillo, H. and Lipman, D., SIAM J. Applied Math. 48:1073, 1988). Methods which are designed to give the largest match between the sequences are generally preferred. Methods to determine identity and similarity are codified in publicly available computer programs including the GCG program package (Devereux J. et al., Nucleic Acids Research 12(1): 387, 1984); BLASTP, BLASTN, and FASTA (Atschul, S.F. et al. J. Molec. Biol. 215: 403-410, 1990). The BLAST X program is publicly available from NCBI and other sources (BLAST

Manual, Altschul, S. et al. NCBI NLM NIH Bethesda, Md. 20894; Altschul, S. et al. J. Mol. Biol. 215: 403-410, 1990).

Isolated nucleic acid molecules encoding a protein having the activity of SEL-1L Protein, and having a sequence which differs from the nucleic acid sequence shown in SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ. ID. NO. 7, or SEQ. ID. NO. 9 due to degeneracy in the genetic code are also within the scope of the invention. Such nucleic acids encode functionally equivalent proteins (e.g., a SEL-1L Protein) but differ in sequence from the sequence of SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ. ID. NO. 7, or SEQ. ID. NO. 9 due to degeneracy in the genetic code. As one example, DNA sequence polymorphisms within the nucleotide sequence of SEL-1L Protein may result in silent mutations which do not affect the amino acid sequence. Variations in one or more nucleotides may exist among individuals within a population due to natural allelic variation. Any and all such nucleic acid variations are within the scope of the invention. DNA sequence polymorphisms may also occur which lead to changes in the amino acid sequence of SEL-1L Protein. These amino acid polymorphisms are also within the scope of the present invention.

Another aspect of the invention provides a nucleic acid molecule which hybridizes under stringent conditions, preferably high stringency conditions to a nucleic acid molecule which comprises a sequence which encodes SEL-1L Protein having the amino acid sequence shown in SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10. Appropriate stringency conditions which promote DNA hybridization are known to those skilled in the art, or can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, 6.0 x sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x SSC at 50°C may be employed. The stringency may be selected based on the conditions used in the wash step. By way of example, the salt concentration in the wash step can be selected from a high stringency of about 0.2 x SSC at 50°C. In addition, the temperature in the wash step can be at high stringency conditions, at about 65°C.

It will be appreciated that the invention includes nucleic acid molecules encoding a SEL-1L Related Protein including truncations of SEL-1L Protein, and analogs of SEL-1L Protein as described herein. It will further be appreciated that variant forms of the nucleic acid molecules of the invention which arise by alternative splicing of an mRNA corresponding to a cDNA of the invention are encompassed by the invention.

An isolated nucleic acid molecule of the invention which comprises DNA can be isolated by preparing a labelled nucleic acid probe based on all or part of the nucleic acid sequence shown in SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ. ID. NO. 7, or SEQ. ID. NO. 9. The labeled nucleic acid probe is used to screen an appropriate DNA library (e.g. a cDNA or genomic DNA library). For example, a cDNA library can be used to isolate a cDNA encoding a SEL-1L Related Protein by screening the library with the labeled probe using standard techniques. Alternatively, a genomic DNA library can be similarly screened to isolate a genomic clone encompassing a gene encoding a SEL-1L Related Protein. Nucleic acids isolated by screening of a cDNA or genomic DNA library can be sequenced by standard techniques.

An isolated nucleic acid molecule of the invention which is DNA can also be isolated by selectively amplifying a nucleic acid encoding a SEL-1L Related Protein using the polymerase chain reaction (PCR) methods and cDNA or genomic DNA. It is possible to design synthetic oligonucleotide primers from the nucleotide sequence shown in SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ.

ID. NO. 7, or SEQ. ID. NO. 9 for use in PCR. A nucleic acid can be amplified from cDNA or genomic DNA using these oligonucleotide primers and standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. cDNA may be prepared from mRNA, by isolating total cellular mRNA by a variety of techniques, for example, by using the guanidinium-thiocyanate extraction procedure of Chirgwin et al., Biochemistry, 18, 5294-5299 (1979). cDNA is then synthesized from the mRNA using reverse transcriptase (for example, Moloney MLV reverse transcriptase available from Gibco/BRL, Bethesda, MD, or AMV reverse transcriptase available from Seikagaku America, Inc., St. Petersburg, FL).

An isolated nucleic acid molecule of the invention which is RNA can be isolated by cloning a cDNA encoding a SEL-1L Related Protein into an appropriate vector which allows for transcription of the cDNA to produce an RNA molecule which encodes a SEL-1L Related Protein. For example, a cDNA can be cloned downstream of a bacteriophage promoter, (e.g. a T7 promoter) in a vector, cDNA can be transcribed *in vitro* with T7 polymerase, and the resultant RNA can be isolated by conventional techniques.

Nucleic acid molecules of the invention may be chemically synthesized using standard techniques. Methods of chemically synthesizing polydeoxynucleotides are known, including but not limited to solid-phase synthesis which, like peptide synthesis, has been fully automated in commercially available DNA synthesizers (See e.g., Itakura et al. U.S. Patent No. 4,598,049; Caruthers et al. U.S. Patent No. 4,458,066; and Itakura U.S. Patent Nos. 4,401,796 and 4,373,071).

Determination of whether a particular nucleic acid molecule encodes a SEL-1L Related Protein can be accomplished by expressing the cDNA in an appropriate host cell by standard techniques, and testing the expressed protein in the methods described herein. A cDNA encoding a SEL-1L Related Protein can be sequenced by standard techniques, such as dideoxynucleotide chain termination or Maxam-Gilbert chemical sequencing, to determine the nucleic acid sequence and the predicted amino acid sequence of the encoded protein.

The initiation codon and untranslated sequences of a SEL-1L Related Protein may be determined using computer software designed for the purpose, such as PC/Gene (IntelliGenetics Inc., Calif.). The intron-exon structure and the transcription regulatory sequences of a gene encoding a SEL-1L Related Protein may be identified by using a nucleic acid molecule of the invention encoding a SEL-1L Related Protein to probe a genomic DNA clone library. Regulatory elements can be identified using standard techniques. The function of the elements can be confirmed by using these elements to express a reporter gene such as the lacZ gene which is operatively linked to the elements. These constructs may be introduced into cultured cells using conventional procedures or into non-human transgenic animal models. In addition to identifying regulatory elements in DNA, such constructs may also be used to identify nuclear proteins interacting with the elements, using techniques known in the art.

In a particular embodiment of the invention, the nucleic acid molecules isolated using the methods described herein are mutant *Sel-1L* gene alleles. The mutant alleles may be isolated from individuals either known or proposed to have a genotype which contributes to the symptoms of Alzheimer's Disease, cancer (e.g. pancreatic cancer), diabetes, or Parkinson's Disease. Mutant alleles and mutant allele products may be used in therapeutic and diagnostic methods described herein. For example, a cDNA of a mutant *Sel-1L* gene may be isolated using PCR as described herein, and the DNA sequence of the mutant allele may be compared to the normal allele to ascertain the mutation(s) responsible for the loss

or alteration of function of the mutant gene product. A genomic library can also be constructed using DNA from an individual suspected of or known to carry a mutant allele, or a cDNA library can be constructed using RNA from tissue known, or suspected to express the mutant allele. A nucleic acid encoding a normal *Sel-1L* gene or any suitable fragment thereof, may then be labeled and used as a probe to identify the corresponding mutant allele in such libraries. Clones containing mutant sequences can be purified and subjected to sequence analysis. In addition, an expression library can be constructed using cDNA from RNA isolated from a tissue of an individual known or suspected to express a mutant *Sel-1L* allele. Gene products made by the putatively mutant tissue may be expressed and screened, for example using antibodies specific for a SEL-1L Related Protein as described herein. Library clones identified using the antibodies can be purified and subjected to sequence analysis.

The sequence of a nucleic acid molecule of the invention, or a fragment of the molecule, may be inverted relative to its normal presentation for transcription to produce an antisense nucleic acid molecule. An antisense nucleic acid molecule may be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art.

2. **Proteins of the Invention**

The SEL-1L protein of the invention is highly expressed in the pancreas, and in the floor plate of the neural tube. The amino acid sequence of SEL-1L Protein comprises the sequence as shown in SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10.

In addition to proteins comprising the amino acid sequence as shown in SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10, the proteins of the present invention include truncations of a SEL-1L Protein, analogs of a SEL-1L Protein, and proteins having sequence identity or similarity to a SEL-1L Protein, and truncations thereof as described herein (i.e. SEL-1L Related Proteins). Truncated proteins may comprise peptides of between 3 and 70 amino acid residues, ranging in size from a tripeptide to a 70 mer polypeptide.

The truncated proteins may have an amino group (-NH2), a hydrophobic group (for example, carbobenzoxyl, dansyl, or T-butyloxycarbonyl), an acetyl group, a 9-fluorenylmethoxy-carbonyl (PMOC) group, or a macromolecule including but not limited to lipid-fatty acid conjugates, polyethylene glycol, or carbohydrates at the amino terminal end. The truncated proteins may have a carboxyl group, an amido group, a T-butyloxycarbonyl group, or a macromolecule including but not limited to lipid-fatty acid conjugates, polyethylene glycol, or carbohydrates at the carboxy terminal end.

The proteins of the invention may also include analogs of SEL-1L Protein, and/or truncations thereof as described herein, which may include, but are not limited to SEL-1L Protein, containing one or more amino acid substitutions, insertions, and/or deletions. Amino acid substitutions may be of a conserved or non-conserved nature. Conserved amino acid substitutions involve replacing one or more amino acids of the SEL-1L Protein amino acid sequence with amino acids of similar charge, size, and/or hydrophobicity characteristics. When only conserved substitutions are made the resulting analog is preferably functionally equivalent to SEL-1L Protein. Non-conserved substitutions involve replacing one or more amino acids of the SEL-1L Protein amino acid sequence with one or more amino acids which possess dissimilar charge, size, and/or hydrophobicity characteristics.

One or more amino acid insertions may be introduced into SEL-1L Protein. Amino acid insertions may consist of single amino acid residues or sequential amino acids ranging from 2 to 15 amino acids in length.

Deletions may consist of the removal of one or more amino acids, or discrete portions from the SEL-1L Protein sequence. The deleted amino acids may or may not be contiguous. The lower limit length of the resulting analog with a deletion mutation is about 10 amino acids, preferably 100 amino acids.

The proteins of the invention include proteins with sequence identity or similarity to SEL-1L Protein and/or truncations thereof as described herein. Such SEL-1L Proteins include proteins whose amino acid sequences are comprised of the amino acid sequences of SEL-1L Protein regions from other species that hybridize under selected hybridization conditions (see discussion of stringent hybridization conditions herein) with a probe used to obtain SEL-1L Protein. These proteins will generally have the same regions which are characteristic of SEL-1L Protein. Preferably a protein will have substantial sequence identity for example, about 50% identity, preferably 70 to 80% identity, more preferably at least 90% to 95% identity, and most preferably 98% identity with the amino acid sequence shown in SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10.

A percent amino acid sequence homology, similarity or identity is calculated as the percentage of aligned amino acids that match the reference sequence using known methods as described herein.

The invention also contemplates isoforms of the proteins of the invention. An isoform contains the same number and kinds of amino acids as the protein of the invention, but the isoform has a different molecular structure. The isoforms contemplated by the present invention preferably have the same properties as the protein of the invention as described herein.

The present invention also includes SEL-1L Related Proteins conjugated with a selected protein, or a marker protein (see below) to produce fusion proteins. Additionally, immunogenic portions of SEL-1L Protein and SEL-1L Protein Related Proteins are within the scope of the invention.

A SEL-1L Related Protein of the invention may be prepared using recombinant DNA methods. Accordingly, the nucleic acid molecules of the present invention having a sequence which encodes a SEL-1L Related Protein of the invention may be incorporated in a known manner into an appropriate expression vector which ensures good expression of the protein. Possible expression vectors include but are not limited to cosmids, plasmids, or modified viruses (e.g. replication defective retroviruses, adenoviruses and adeno-associated viruses), so long as the vector is compatible with the host cell used.

The invention therefore contemplates a recombinant expression vector of the invention containing a nucleic acid molecule of the invention, and the necessary regulatory sequences for the transcription and translation of the inserted protein-sequence. Suitable regulatory sequences may be derived from a variety of sources, including bacterial, fungal, viral, mammalian, or insect genes (For example, see the regulatory sequences described in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Selection of appropriate regulatory sequences is dependent on the host cell chosen as discussed below, and may be readily accomplished by one of ordinary skill in the art. The necessary regulatory sequences may be supplied by the native SEL-1L Protein and/or its flanking regions.

The invention further provides a recombinant expression vector comprising a DNA nucleic acid molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is linked to a regulatory sequence in a manner which allows for expression, by transcription of

the DNA molecule, of an RNA molecule which is antisense to the nucleic acid sequence of SEQ. ID. NO. 1; SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ. ID. NO. 7, or SEQ. ID. NO. 9 or a fragment thereof. Regulatory sequences linked to the antisense nucleic acid can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance a viral promoter and/or enhancer, or regulatory sequences can be chosen which direct tissue or cell type specific expression of antisense RNA.

The recombinant expression vectors of the invention may also contain a marker gene which facilitates the selection of host cells transformed or transfected with a recombinant molecule of the invention. Examples of marker genes are genes encoding a protein such as G418 and hygromycin which confer resistance to certain drugs, β-galactosidase, chloramphenicol acetyltransferase, firefly luciferase, or an immunoglobulin or portion thereof such as the Fc portion of an immunoglobulin preferably IgG. The markers can be introduced on a separate vector from the nucleic acid of interest.

The recombinant expression vectors may also contain genes which encode a fusion moiety which provides increased expression of the recombinant protein; increased solubility of the recombinant protein; and aid in the purification of the target recombinant protein by acting as a ligand in affinity purification. For example, a proteolytic cleavage site may be added to the target recombinant protein to allow separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Typical fusion expression vectors include pGEX (Amrad Corp., Melbourne, Australia), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the recombinant protein.

The recombinant expression vectors may be introduced into host cells to produce a transformant host cell. "Transformant host cells" include host cells which have been transformed or transfected with a recombinant expression vector of the invention. The terms "transformed with", "transfected with", "transformation" and "transfection" encompass the introduction of a nucleic acid (e.g. a vector) into a cell by one of many standard techniques. Prokaryotic cells can be transformed with a nucleic acid by, for example, electroporation or calcium-chloride mediated transformation. Nucleic acid can be introduced into mammalian cells via conventional techniques such as calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofectin, electroporation or microinjection. Suitable methods for transforming and transfecting host cells can be found in Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory press (1989)), and other laboratory textbooks.

Suitable host cells include a wide variety of prokaryotic and eukaryotic host cells. For example, the proteins of the invention may be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus), yeast cells or mammalian cells. Other suitable host cells can be found in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1991).

A host cell may also be chosen which modulates the expression of an inserted nucleic acid sequence, or modifies (e.g. glycosylation or phosphorylation) and processes (e.g. cleaves) the protein in a desired fashion. Host systems or cell lines may be selected which have specific and characteristic mechanisms for post-translational processing and modification of proteins. For example, eukaryotic host cells including CHO, VERO, BHK, HeLA, COS, MDCK, 293, 3T3, and WI38 may be used. For long-term high-yield stable expression of the protein, cell lines and host systems which stably express the gene product may be engineered.

Host cells and in particular cell lines produced using the methods described herein may be particularly useful in screening and evaluating compounds that modulate the activity of a SEL-1L Related Protein.

The proteins of the invention may also be expressed in non-human transgenic animals including but not limited to mice, rats, rabbits, guinea pigs, micro-pigs, goats, sheep, pigs, non-human primates (e.g. baboons, monkeys, and chimpanzees) [see Hammer et al. (Nature 315:680-683, 1985), Palmiter et al. (Science 222:809-814, 1983), Brinster et al. (Proc Natl. Acad. Sci USA 82:44384442, 1985), Palmiter and Brinster (Cell. 41:343-345, 1985) and U.S. Patent No. 4,736,866)]. Procedures known in the art may be used to introduce a nucleic acid molecule of the invention encoding a SEL-1L Related Protein into animals to produce the founder lines of transgenic animals. Such procedures include pronuclear microinjection, retrovirus mediated gene transfer into germ lines, gene targeting in embryonic stem cells, electroporation of embryos, and sperm-mediated gene transfer.

The present invention contemplates a transgenic animal that carries the *Sel-1L* gene in all their cells, and animals which carry the transgene in some but not all their cells. The transgene may be integrated as a single transgene or in concatamers. The transgene may be selectively introduced into and activated in specific cell types (See for example, Lasko et al, 1992 Proc. Natl. Acad. Sci. USA 89: 6236). The transgene may be integrated into the chromosomal site of the endogenous gene by gene targeting. The transgene may be selectively introduced into a particular cell type inactivating the endogenous gene in that cell type (See Gu et al Science 265: 103-106).

The expression of a recombinant SEL-1L Related Protein in a transgenic animal may be assayed using standard techniques. Initial screening may be conducted by Southern Blot analysis, or PCR methods to analyze whether the transgene has been integrated. The level of mRNA expression in the tissues of transgenic animals may also be assessed using techniques including Northern blot analysis of tissue samples, *in situ* hybridization, and RT-PCR. Tissue may also be evaluated immunocytochemically using antibodies against SEL-1L Protein.

Proteins of the invention may also be prepared by chemical synthesis using techniques well known in the chemistry of proteins such as solid phase synthesis (Merrifield, 1964, J. Am. Chem. Assoc. 85:2149-2154) or synthesis in homogenous solution (Houbenweyl, 1987, Methods of Organic Chemistry, ed. E. Wansch, Vol. 15 I and II, Thieme, Stuttgart).

N-terminal or C-terminal fusion proteins comprising a SEL-1L Related Protein of the invention conjugated with other molecules, such as proteins, may be prepared by fusing, through recombinant techniques, the N-terminal or C-terminal of a SEL-1L Related Protein, and the sequence of a selected protein or marker protein with a desired biological function. The resultant fusion proteins contain SEL-1L Protein fused to the selected protein or marker protein as described herein. Examples of proteins which may be used to prepare fusion proteins include immunoglobulins, glutathione-S-transferase (GST), hemagglutinin (HA), and truncated myc.

3. **Antibodies**

SEL-1L Related Proteins of the invention can be used to prepare antibodies specific for the proteins. Antibodies can be prepared which bind a distinct epitope in an unconserved region of the protein. An unconserved region of the protein is one that does not have substantial sequence homology to other proteins. A region from a conserved region such as a well-characterized domain can also be used to

prepare an antibody to a conserved region of a SEL-1L Related Protein. For example, antibodies may be prepared against the Sel-1L PTB domain binding site. Antibodies having specificity for a SEL-1L Related Protein may also be raised from fusion proteins created by expressing fusion proteins in bacteria as described herein.

The invention can employ intact monoclonal or polyclonal antibodies, and immunologically active fragments (e.g. a Fab, $(Fab)_2$ fragment, or Fab expression library fragments and epitope-binding fragments thereof), an antibody heavy chain, and antibody light chain, a genetically engineered single chain Fv molecule (Ladner et al, U.S. Pat. No. 4,946,778), or a chimeric antibody, for example, an antibody which contains the binding specificity of a murine antibody, but in which the remaining portions are of human origin. Antibodies including monoclonal and polyclonal antibodies, fragments and chimeras, may be prepared using methods known to those skilled in the art.

4. Applications of the Nucleic Acid Molecules, SEL-1L Related Proteins, and Antibodies of the Invention

The nucleic acid molecules, SEL-1L Related Proteins, and antibodies of the invention may be used in the prognostic and diagnostic evaluation of Alzheimer's Disease, Insulin Dependent Diabetes Mellitus (IDDM), cancer (e.g. pancreatic cancer), stroke, vascular dementia, neurodegenerative diseases such as Parkinson's Disease, and conditions requiring modulation of fat or cholesterol metabolism such as coronary heart disease, and the identification of subjects with a predisposition to such conditions (Section 4.1.1 and 4.1.2). Methods for detecting nucleic acid molecules and SEL-1L Related Proteins of the invention, can be used to monitor Alzheimer's Disease, Insulin Dependent Diabetes Mellitus, cancer (e.g. pancreatic cancer), stroke, vascular dementia, neurodegenerative diseases such as Parkinson's Disease, and conditions requiring modulation of fat or cholesterol, by detecting and localizing SEL-1L Related Proteins and nucleic acid molecules encoding SEL-1L Related Proteins. It would also be apparent to one skilled in the art that the methods described herein may be used to study the developmental expression of SEL-1L Related Proteins and, accordingly, will provide further insight into the role of SEL-1L Related Proteins. The applications of the present invention also include methods for the identification of compounds that modulate the biological activity of *Sel-1L* or SEL-1L Related Proteins (Section 4.2). The compounds, antibodies etc. may be used for the treatment of Alzheimer's Disease, diabetes, cancer (e.g. pancreatic cancer), stroke, vascular dementia, neurodegenerative diseases such as Parkinson's Disease, and conditions requiring modulation of fat or cholesterol metabolism such as coronary heart disease (Section 4.4).

4.1 Diagnostic Methods

A variety of methods can be employed for the diagnostic and prognostic evaluation of Alzheimer's Disease, IDDM, cancer (e.g. pancreatic cancer), vascular dementia, stroke, neurodegenerative diseases such as Parkinson's Disease, and conditions requiring modulation of fat or cholesterol metabolism such as coronary heart disease, and the identification of subjects with a predisposition to such conditions. Such methods may, for example, utilize nucleic acid molecules of the invention, and fragments thereof, and antibodies directed against SEL-1L Related Proteins, including peptide fragments. In particular, the nucleic acids and antibodies may be used, for example, for: (1) the detection of the presence of *Sel-1L* mutations, or the detection of either over- or under-expression of *Sel-1L* mRNA relative to a non-disorder state or the qualitative or quantitative detection of alternatively spliced forms of *Sel-1L* transcripts which may correlate with certain conditions or susceptibility toward such conditions; and (2) the detection of

either an over- or an under-abundance of SEL-1L Related Proteins relative to a non- disorder state or the presence of a modified (e.g., less than full length) SEL-1L Protein which correlates with a disorder state, or a progression toward a disorder state.

The methods described herein may be performed by utilizing pre-packaged diagnostic kits comprising at least one specific *Sel-1L* nucleic acid or antibody described herein, which may be conveniently used, e.g., in clinical settings, to screen and diagnose patients and to screen and identify those individuals exhibiting a predisposition to developing a disorder.

Nucleic acid-based detection techniques are described, below, in Section 4.1.1. Peptide detection techniques are described, below, in Section 4.1.2. The samples that may be analyzed using the methods of the invention include those which are known or suspected to express *Sel-1L* or contain SEL-1L Related Proteins. The samples may be derived from a patient or a cell culture, and include but are not limited to biological fluids, tissue extracts, freshly harvested cells, and lysates of cells which have been incubated in cell cultures.

4.1.1 Methods for Detecting Nucleic Acid Molecules of the Invention

The nucleic acid molecules of the invention allow those skilled in the art to construct nucleotide probes for use in the detection of nucleic acid sequences of the invention in samples. Suitable probes include nucleic acid molecules based on nucleic acid sequences encoding at least 5 sequential amino acids from regions of the SEL-1L Protein (see SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10), preferably they comprise 15 to 30 nucleotides. A nucleotide probe may be labeled with a detectable substance such as a radioactive label which provides for an adequate signal and has sufficient half-life such as ^{32}P, ^{3}H, ^{14}C or the like. Other detectable substances which may be used include antigens that are recognized by a specific labeled antibody, fluorescent compounds, enzymes, antibodies specific for a labeled antigen, and luminescent compounds. An appropriate label may be selected having regard to the rate of hybridization and binding of the probe to the nucleotide to be detected and the amount of nucleotide available for hybridization. Labeled probes may be hybridized to nucleic acids on solid supports such as nitrocellulose filters or nylon membranes as generally described in Sambrook et al, 1989, Molecular Cloning, A Laboratory Manual (2nd ed.). The nucleic acid probes may be used to detect genes, preferably in human cells, that encode SEL-1L Related Proteins. The nucleotide probes may also be useful in the diagnosis of Alzheimer's Disease, IDDM, cancer (e.g. pancreatic cancer), stroke, vascular dementia, neurodegenerative diseases such as Parkinson's Disease, and conditions requiring modulation of fat or cholesterol metabolism such as coronary heart disease; in monitoring the progression of these conditions; or monitoring a therapeutic treatment.

The probe may be used in hybridization techniques to detect genes that encode Sel-1 Related Proteins. The technique generally involves contacting and incubating nucleic acids (e.g. recombinant DNA molecules, cloned genes) obtained from a sample from a patient or other cellular source with a probe of the present invention under conditions favorable for the specific annealing of the probes to complementary sequences in the nucleic acids. After incubation, the non-annealed nucleic acids are removed, and the presence of nucleic acids that have hybridized to the probe if any are detected.

The detection of nucleic acid molecules of the invention may involve the amplification of specific gene sequences using an amplification method such as PCR, followed by the analysis of the amplified

molecules using techniques known to those skilled in the art. Suitable primers can be routinely designed by one of skill in the art.

Genomic DNA may be used in hybridization or amplification assays of biological samples to detect abnormalities involving *Sel-1L* structure, including point mutations, insertions, deletions, and chromosomal rearrangements. For example, direct sequencing, single stranded conformational polymorphism analyses, heteroduplex analysis, denaturing gradient gel electrophoresis, chemical mismatch cleavage, and oligonucleotide hybridization may be utilized.

Genotyping techniques known to one skilled in the art can be used to type polymorphisms that are in close proximity to the mutations in the *Sel-1L* gene. The polymorphisms may be used to identify individuals in families that are likely to carry mutations. If a polymorphism exhibits linkage disequalibrium with mutations in the *Sel-1L* gene, it can also be used to screen for individuals in the general population likely to carry mutations. Polymorphisms which may be used include restriction fragment length polymorphisms (RFLPs), single-base polymorphisms, and simple sequence repeat polymorphisms (SSLPs).

A probe of the invention may be used to directly identify RFLPs. A probe or primer of the invention can additionally be used to isolate genomic clones such as YACs, BACs, PACs, cosmids, phage or plasmids. The DNA in the clones can be screened for SSLPs using hybridization or sequencing procedures.

Hybridization and amplification techniques described herein may be used to assay qualitative and quantitative aspects of *Sel-1L* expression. For example, RNA may be isolated from a cell type or tissue known to express *Sel-1L* (e.g. pancreas) and tested utilizing the hybridization (e.g. standard Northern analyses) or PCR techniques referred to herein. The techniques may be used to detect differences in transcript size which may be due to normal or abnormal alternative splicing. The techniques may be used to detect quantitative differences between levels of full length and/or alternatively splice transcripts detected in normal individuals relative to those individuals exhibiting symptoms of Alzheimer's Disease or other disease conditions.

The primers and probes may be used in the above described methods *in situ* i.e directly on tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections.

4.1.2 Methods for Detecting SEL-1L Related Proteins

Antibodies specifically reactive with a SEL-1L Related Protein, or derivatives, such as enzyme conjugates or labeled derivatives, may be used to detect SEL-1L Related Proteins in various samples (e.g. biological materials). They may be used as diagnostic or prognostic reagents and they may be used to detect abnormalities in the level of SEL-1L Related Proteins expression, or abnormalities in the structure, and/or temporal, tissue, cellular, or subcellular location of a SEL-1L Related Protein. Antibodies may also be used to screen potentially therapeutic compounds *in vitro* to determine their effects on Alzheimer's Disease, and other conditions. *In vitro* immunoassays may also be used to assess or monitor the efficacy of particular therapies. The antibodies of the invention may also be used *in vitro* to determine the level of *Sel-1L* expression in cells genetically engineered to produce a SEL-1L Related Protein.

The antibodies may be used in any known immunoassays which rely on the binding interaction between an antigenic determinant of a SEL-1L Related Protein and the antibodies. Examples of such assays are radioimmunoassays, enzyme immunoassays (e.g. ELISA), immunofluorescence,

immunoprecipitation, latex agglutination, hemagglutination, and histochemical tests. The antibodies may be used to detect and quantify SEL-1L Related Proteins in a sample in order to determine its role in particular cellular events or pathological states, and to diagnose and treat such pathological states.

In particular, the antibodies of the invention may be used in immuno-histochemical analyses, for example, at the cellular and sub-subcellular level, to detect SEL-1L Related Protein, to localize it to particular cells and tissues, and to specific subcellular locations, and to quantitate the level of expression.

Cytochemical techniques known in the art for localizing antigens using light and electron microscopy may be used to detect a SEL-1L Related Protein. Generally, an antibody of the invention may be labeled with a detectable substance and SEL-1L Related Protein may be localised in tissues and cells based upon the presence of the detectable substance. Examples of detectable substances include, but are not limited to, the following: radioisotopes (e.g., ^{3}H, ^{14}C, ^{35}S, ^{125}I, ^{131}I), fluorescent labels (e.g., FITC, rhodamine, lanthanide phosphors), luminescent labels such as luminol; enzymatic labels (e.g., horseradish peroxidase, beta-galactosidase, luciferase, alkaline phosphatase, acetylcholinesterase), biotinyl groups (which can be detected by marked avidin e.g., streptavidin containing a fluorescent marker or enzymatic activity that can be detected by optical or calorimetric methods), predetermined polypeptide epitopes recognized by a secondary reporter (e.g., leucine zipper pair sequences, binding sites for secondary antibodies, metal binding domains, epitope tags). In some embodiments, labels are attached via spacer arms of various lengths to reduce potential steric hindrance. Antibodies may also be coupled to electron dense substances, such as ferritin or colloidal gold, which are readily visualised by electron microscopy.

The antibody or sample may be immobilized on a carrier or solid support which is capable of immobilizing cells, antibodies etc. For example, the carrier or support may be nitrocellulose, or glass, polyacrylamides, gabbros, and magnetite. The support material may have any possible configuration including spherical (e.g. bead), cylindrical (e.g. inside surface of a test tube or well, or the external surface of a rod), or flat (e.g. sheet, test strip). Indirect methods may also be employed in which the primary antigen-antibody reaction is amplified by the introduction of a second antibody, having specificity for the antibody reactive against SEL-1L Related Protein. By way of example, if the antibody having specificity against a SEL-1L Related Protein is a rabbit IgG antibody, the second antibody may be goat anti-rabbit gamma-globulin labeled with a detectable substance as described herein.

Where a radioactive label is used as a detectable substance, a SEL-1L Related Protein may be localized by radioautography. The results of radioautography may be quantitated by determining the density of particles in the radioautographs by various optical methods, or by counting the grains.

4.2 Methods for Identifying or Evaluating Substances/Compounds

The methods described herein are designed to identify substances that modulate the biological activity of a SEL-1L Related Protein including substances that bind to SEL-1L Related Proteins, or bind to other proteins that interact with a SEL-1L Related Protein, to compounds that interfere with, or enhance the interaction of a SEL-1L Related Protein and substances that bind to the SEL-1L Related Protein or other proteins that interact with a SEL-1L Related Protein. Methods are also utilized that identify compounds that bind to *Sel-1L* regulatory sequences.

The substances and compounds identified using the methods of the invention include but are not limited to peptides such as soluble peptides including Ig-tailed fusion peptides, members of random peptide libraries and combinatorial chemistry-derived molecular libraries made of D- and/or L-

configuration amino acids, phosphopeptides (including members of random or partially degenerate, directed phosphopeptide libraries), antibodies [e.g. polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, single chain antibodies, fragments, (e.g. Fab, $F(ab)_2$, and Fab expression library fragments, and epitope-binding fragments thereof)], and small organic or inorganic molecules. The substance or compound may be an endogenous physiological compound or it may be a natural or synthetic compound.

Substances which modulate a SEL-1L Related Protein can be identified based on their ability to bind to a SEL-1L Related Protein. Therefore, the invention also provides methods for identifying substances which bind to a SEL-1L Related Protein. Substances identified using the methods of the invention may be isolated, cloned and sequenced using conventional techniques.

Substances which can bind with a SEL-1L Related Protein may be identified by reacting a SEL-1L Related Protein with a test substance which potentially binds to a SEL-1L Related Protein, under conditions which permit the formation of substance-SEL-1L Related Protein complexes and removing and/or detecting the complexes. The detection of complexes indicates the test substance binds to the SEL-1L Related Protein. The complexes can be detected by assaying for substance-SEL-1L Related Protein complexes, for free substance, or for non-complexed SEL-1L Related Protein. Conditions which permit the formation of substance-SEL-1L Related Protein complexes may be selected having regard to factors such as the nature and amounts of the substance and the protein.

The substance-protein complex, free substance or non-complexed proteins may be isolated by conventional isolation techniques, for example, salting out, chromatography, electrophoresis, gel filtration, fractionation, absorption, polyacrylamide gel electrophoresis, agglutination, or combinations thereof. To facilitate the assay of the components, antibody against SEL-1L Related Protein or the substance, or labeled SEL-1L Related Protein, or a labeled substance may be utilized. The antibodies, proteins, or substances may be labeled with a detectable substance as described above.

A SEL-1L Related Protein, or the substance used in the method of the invention may be insolubilized. For example, a SEL-1L Related Protein, or substance may be bound to a suitable carrier such as agarose, cellulose, dextran, Sephadex, Sepharose, carboxymethyl cellulose polystyrene, filter paper, ion-exchange resin, plastic film, plastic tube, glass beads, polyamine-methyl vinyl-ether-maleic acid copolymer, amino acid copolymer, ethylene-maleic acid copolymer, nylon, silk, etc. The carrier may be in the shape of, for example, a tube, test plate, beads, disc, sphere etc. The insolubilized protein or substance may be prepared by reacting the material with a suitable insoluble carrier using known chemical or physical methods, for example, cyanogen bromide coupling.

The invention also contemplates a method for evaluating a compound for its ability to modulate the biological activity of a SEL-1L Related Protein of the invention, by assaying for an agonist or antagonist (i.e. enhancer or inhibitor) of the binding of a SEL-1L Related Protein with a substance which binds with a SEL-1L Related Protein.

The term "agonist" refers to a molecule which increases the amount of, or prolongs the duration of, the activity of a complex of a SEL-1L Related Protein and substance which binds with the SEL-1L Related Protein, or molecules of the complex, or which increases the formation of such complexes thereby increasing the amount of, or prolongs the duration of, the activity of the complex or molecules of the complex. The stimulation may be direct, or indirect, or by a competitive or non-competitive mechanism.

The term "antagonist", as used herein, refers to a molecule which decreases the amount of, or shortens the duration of, the activity of a complex of a SEL-1L Related Protein and substance which binds with the SEL-1L Related Protein, or molecules of the complex, or decreases the formation of such complexes thereby decreasing the amount of, or shortens the duration of, the activity of the complex or molecules of the complex. The inhibition may be direct, or indirect, or by a competitive or non-competitive mechanism

The basic method for evaluating if a compound is an agonist or antagonist of the binding of a SEL-1L Related Protein and a substance that binds to the protein, is to prepare a reaction mixture containing the SEL-1L Related Protein and the substance under conditions which permit the formation of substance-SEL-1L Related Protein complexes, in the presence of a test compound. The test compound may be initially added to the mixture, or may be added subsequent to the addition of the SEL-1L Related Protein and substance. Control reaction mixtures without the test compound or with a placebo are also prepared. The formation of complexes is detected and the formation of complexes in the control reaction but not in the reaction mixture indicates that the test compound interferes with the interaction of the SEL-1L Related Protein and substance. Increased complex formation indicates that the test compound enhances the interaction. The reactions may be carried out in the liquid phase or the SEL-1L Related Protein, substance, or test compound may be immobilized as described herein. The ability of a compound to modulate the biological activity of a Sel-1L Related Protein of the invention may also be tested by measuring activation of the protein (e.g. by assaying for phosphorylation of tyrosine residues using conventional methods), or determining the biological effects on cells such as inhibition or stimulation of proliferation or differentiation of cells, degradation of proteins in the cells, or migration of the cells.

It will be understood that the agonists and antagonists i.e. inhibitors and enhancers that can be assayed using the methods of the invention may act on one or more of the binding sites on the protein or substance including agonist binding sites, competitive antagonist binding sites, non-competitive antagonist binding sites or allosteric sites.

The invention also makes it possible to screen for antagonists that inhibit the effects of an agonist of the interaction of SEL-1L Related Protein with a substance which is capable of binding to the SEL-1L Related Protein. Thus, the invention may be used to assay for a compound that competes for the same binding site of a SEL-1L Related Protein.

The invention also contemplates methods for identifying compounds that bind to proteins that interact with a SEL-1L Related Protein. Protein-protein interactions may be identified using conventional methods such as co-immunoprecipitation, crosslinking and co-purification through gradients or chromatographic columns. Methods may also be employed that result in the simultaneous identification of genes which encode proteins interacting with a SEL-1L Related Protein. These methods include probing expression libraries with labeled SEL-1L Related Protein.

Two-hybrid systems may also be used to detect protein interactions *in vivo*. Generally, plasmids are constructed that encode two hybrid proteins. A first hybrid protein consists of the DNA-binding domain of a transcription activator protein fused to a SEL-1L Related Protein, and the second hybrid protein consists of the transcription activator protein's activator domain fused to an unknown protein encoded by a cDNA which has been recombined into the plasmid as part of a cDNA library. The plasmids are transformed into a strain of yeast (e.g. *S. cerevisiae*) that contains a reporter gene (e.g. lacZ, luciferase,

alkaline phosphatase, horseradish peroxidase) whose regulatory region contains the transcription activator's binding site. The hybrid proteins alone cannot activate the transcription of the reporter gene. However, interaction of the two hybrid proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

It will be appreciated that fusion proteins may be used in the above-described methods. In particular, SEL-1L Related Proteins fused to a glutathione-S-transferase may be used in the methods.

The reagents suitable for applying the methods of the invention to evaluate compounds that modulate a SEL-1L Related Protein may be packaged into convenient kits providing the necessary materials packaged into suitable containers. The kits may also include suitable supports useful in performing the methods of the invention.

4.3 Screening Methods and Peptides Based on the PTB Domain Binding Site and PDZ Domains

SEL-1L contains a PTB domain binding site and therefore substances and methods are provided for modulating the interaction of a PTB domain containing protein and a SEL-1L PTB domain binding site. SEL-1L also contains three PDZ domains (consensus sequences associated with the domains are GLGF at position #231, GLGQ at position #349, and GLGM at position #423, and having the characteristic PDZ 2° structures) and therefore substances and methods are provided for modulating the interaction of a SEL-1L PDZ domain and a protein containing a PDZ domain binding motif.

The term "PTB domain containing protein" refers to a protein or peptide, or part thereof which comprises or consists of a PTB domain. A PTB domain is a region of ~160 amino acids which was originally identified in Shc and Sck (Kavanaugh, V.M. Et al., 1995 Science, 268:1177-1179; Bork, RP, and Margolis, B, Cell, Vol 80:693-694, 1995; Craparo, A., et al., 1995, J. Biol. Chem. 270:15639-15643; van der Geer, P., & Pawson, T., 1995, TIBS 20:277-280; Batzer, A.G., et al., Mol. Cell. Biol. 1995, 15:4403-4409; and Trub, T., et al., 1995, J. Biol. Chem. 270:18205-18208; van der Geer et al., Current Biology 5(4):404, 1995)). The PTB domain comprises residues 46 to 208 in the 52 kDa isoform of Shc. Examples of PTB domain containing proteins are Shc and Sck, IRS-1, IRS-2, and NUMB and homologues of these proteins. Other proteins that contain homologous PTB domains have been identified using data base search methods (Bork, RP, and Margolis, B. Cell, Vol 80:693-694, 1995). PTB domain containing proteins may also be identified by screening a cDNA expression library with a protein containing a sequence with high affinity to PTB domains, i.e. a PTB domain binding sequence or a peptide of the invention which may be labeled. PTB domain containing proteins may also be screened using antibodies specific for the PTB domain. For example, a PTB domain that binds to the consensus sequence Asn-Leu-Gly-Tyr-Met-His-Glu may be identified by screening a cDNA expression library with proteins based on the consensus sequence. One could use PCR (Wilks, A.F., Proc. Natl. Acad. Sci. U.S.A. Vol. 86, pp. 1603-1607, March 1989) or low stringency screening (Hanks, S.K., Proc. Natl. Acad. Sci. U.S.A. Vol. 84, pp 388-392, January 1987) with the PTB domain specific probe.

The term "SEL-1L PTB domain binding site" refers to a sequence in SEL-1L, and highly conserved in hSEL-1L, mSEL-1L and Hrd3p with affinity to PTB domains. The PTB domain binding site is marked by an asterisk in Figure 2B. The methods of the invention may use the complete SEL-1L protein or a fragment thereof containing the PTB domain binding site. The site comprises the sequence Asn-Leu-Gly-Tyr-Met-His-Glu.

The term "SEL-1L PDZ domain" refers to a protein or peptide, or part thereof, which comprises or consists of one or more of the following PDZ consensus motifs: GLGF at position #231, GLGQ at

position #349, and GLGM at position #423. The general characterisitics of PDZ domains are described in the Structural Classification of Proteins (scop) database. The PDZ domain is about 70 amino acids, starting about 3 amino acids upstream of the motif GLGX and it adopts certain characteristic 2° structures.

The term "protein containing a PDZ domain binding motif" refers to a protein or peptide, or part thereof which comprises or consists of a sequence with affinity to a SEL-1L PDZ domain. Examples of such proteins include the NOTCH receptor, and a portion thereof that contains the PDZ binding motif.

The phrase "modulate the interaction of" refers to the ability of a substance to modulate (i.e. inhibit or stimulate) the binding of a PTB domain containing protein to a SEL-1L PTB domain binding site or a SEL-1L PDZ domain and a protein containing a PDZ domain binding motif, thereby affecting regulatory pathways that control gene expression, cell division, cytoskeletal architecture and cell metabolism. Examples of such regulatory pathways are the Ras pathway, the pathway that regulates the breakdown of polyphosphoinositides through phospholipase C, and PI-3-kinase activated pathways, such as the rapamycin-sensitive protein kinase B (PKB/Akt) pathway.

By being "derived from" a domain or binding motif or site is meant any molecular entity which is identical or substantially equivalent to the native domain or binding motif of a molecule. A peptide derived from a specific domain or binding motif may encompass the amino acid sequence of a naturally occurring domain or binding motif, any portion thereof, or other molecular entity that functions to bind to an associated molecule. A peptide derived from such a domain or binding motif will interact directly or indirectly with an associated molecule in such a way as to mimic the native domain or binding motif. Such peptides may include competitive inhibitors, peptide mimetics, and the like.

The invention provides a method for screening for a substance that modulates the interaction of a SEl-1L PTB domain binding site and a PTB domain containing protein comprising

a. reacting a SEL-1L PTB domain binding site, a PTB domain containing protein, and a test substance, under conditions which permit the binding of the SEL-1L PTB domain binding site and the PTB domain containing protein; and
b. assaying for inhibition or stimulation of the binding of the SEL-1L PTB domain binding site and the PTB domain containing protein

wherein detection of inhibition or stimulation of binding relative to binding in the absence of the test substance indicates that the substance modulates the interaction of a SEL-1L PTB domain binding site and a PTB domain containing protein.

The invention also provides a method for screening for a substance that modulates the interaction of a SEL-1L PDZ domain and a protein containing a PDZ domain binding motif comprising

a. reacting a SEL-1L PDZ domain, a protein containing a PDZ domain binding motif, and a test substance, under conditions which permit the binding of the SEL-1L PDZ domain and the protein containing a PDZ domain binding motif; and
b. assaying for inhibition or stimulation of the binding of the SEL-1L PDZ domain and the protein containing a PDZ domain binding motif;

wherein detection of inhibition or stimulation of binding relative to binding in the absence of the test substance indicates that the substance modulates the interaction a SEL-1L PDZ domain and a protein containing a PDZ domain binding motif.

The inhibition or stimulation of the binding of the SEL-1L PTB domain binding site and the PTB domain containing protein may be assayed by measuring complexes of the SEL-1L PTB domain binding site and the PTB domain containing protein, free SEL-1L PTB domain binding site, and/or free PTB domain containing proteins; or by measuring activation of the PTB domain containing protein or a protein containing the SEL-1L PTB domain binding site. The inhibition or stimulation of the binding of the SEL-1L PDZ domain and the protein containing a PDZ domain binding motif, may be assayed by measuring complexes of the SEL-1L PDZ domain and the protein containing a PDZ domain binding motif; free SEL-1L PDZ domain, and/or free proteins containing a PDZ domain binding motif; or by measuring activation of the protein containing a PDZ domain binding motif or SEL-1L. Complexes and free proteins may be isolated by conventional isolation techniques such as salting out, chromatography, electrophoresis, gel filtration, fractionation, absorption, polyacrylamide gel electrophoresis, agglutination, or combinations thereof. To facilitate the assay of components, antibody against the substance, PTB domain binding site, PTB containing protein, SEL-1L PDZ domain, or protein containing a PDZ domain binding motif, or a labeled substance, PTB domain binding site, PTB containing protein, SEL-1L PDZ domain, or the protein containing a PDZ domain binding motif may be utilized. In addition, any of the components of the assay may be insolubilized or expressed on the surface of a cell.

The activation of the PTB domain containing protein, a protein containing the SEL-1L PTB domain binding site, protein containing a PDZ domain binding motif, or SEL-1L may be measured by assaying for phosphorylation of the tyrosine residues, using known techniques such as those using anti-phosphotyrosine antibodies and labeled phosphorous. For example, immunoblots of the complexes may be analyzed by autoradiography (^{32}P samples) or may be blocked and probed with anti-phosphotyrosine antibodies as described in Koch et al, 1989 (Mol. Cell. Biol. 9:4131-4141). The method of the invention may use cells, and inhibition or stimulation may be assayed by determining a biological effect on the cells such as inhibition or stimulation of proliferation or differentiation of the cells, degradation of proteins in the cells, or migration of the cells.

The invention also provides peptide molecules which bind to and inhibit the interactions of a Sel-1L PTB domain binding site and a PTB domain containing protein. The peptide molecules are derived from the SEL-1L PTB domain binding site. For example, peptides of the invention comprise amino acids 641 to 647 of SEL-1L (i.e. Asn-Leu-Gly-Tyr-Met-His-Glu or NLGYMHE). Other proteins containing this binding site sequence may be identified with a protein homology or identity search, for example by searching available databases such as GenBank or SwissProt and various search algorithms and/or programs may be used including FASTA, BLAST (available as a part of the GCG sequence analysis package, University of Wisconsin, Madison, Wis.), or ENTREZ (National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD).

In accordance with an aspect of the invention a peptide of the formula I is provided

$$X_n\text{- Asn-Leu-Gly-Tyr-Met-His-Glu -}Y_n \qquad \text{I}$$

wherein X and Y represent any amino acid and n is 0 to 50, preferably 0 to 20, more preferably 0 to 10. In a most preferred embodiment of the invention the amino acids for X and Y are selected so that they correspond to the sequences flanking NLGYMHE of the SEL-1L sequence.

Preferred peptides of the invention include: Asn-Leu-Gly-Tyr, Asn-Leu-Gly-Tyr-Met, Asn-Leu-Gly-Tyr-Met-His-Glu; Asn-Leu-Gly-Tyr-Met-His-Glu-Lys; Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly;

Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu; Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly; Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile; Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile-Lys; Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile-Lys-Gln; Phe-Asn-Leu-Gly-Tyr-Met-His-Glu; Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu; Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu; Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu; Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu; Ser-Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu; His-Ser-Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu; Ser-Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys; Ser-Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly; ; Ser-Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu; Ser-Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly; Ser-Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile; Ser-Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile-Lys; Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile-Lys; Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile-Lys;Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile-Lys; Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile-Lys; Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile-Lys-Gln; Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile-Lys; Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile; His-Ser-Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile; Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile; ; Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly; Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu; Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly; Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu; Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu; Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly; Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly; Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu; Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly; Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu; Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys; Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys; Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys; Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys; Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu; Phe-Asn-Leu-Gly-Tyr-Met-His-Glu; and Asp-Val-Asp-Tyr-Glu-Thr-Ala-Phe-Ile-His-Tyr-Arg-Leu-Ala-Ser-Glu-Gln-Gln-His-Ser-Ala-Gln-Ala-Met-Phe-Asn-Leu-Gly-Tyr-Met-His-Glu-Lys-Gly-Leu-Gly-Ile-Lys-Gln-Asp-Ile-His-Leu-Ala-Lys-Arg-Phe-Tyr.

The invention also provides peptide molecules which bind to and inhibit the interactions of a SEL-1L PDZ domain and a protein containing a PDZ domain binding motif. The peptide molecules are derived from a SEL-1L PDZ domain, or a binding site on a protein containing a PDZ domain binding motif. For example, peptides of the invention comprise the amino acids GLGF at position #231, GLGQ at position #349, and GLGM at position #423, and optionally the sequences flanking these sequences in the SEL-1L protein. The peptide molecules can be derived from the PDZ binding motif at the C-terminus of the NOTCH receptor which has been shown to associate with Dishevelled. Other proteins containing a PDZ domain binding motif sequence may be identified with a protein homology or identity search, for example by searching available databases such as GenBank or SwissProt and various search algorithms and/or programs may be used including FASTA, BLAST (available as a part of the GCG sequence analysis package, University of Wisconsin, Madison, Wis.), or ENTREZ (National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD).

Examples of peptides derived from a Sel-1 PDZ domain include Thr-Glu-Glu-Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu; Glu-Glu-Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu; Glu-Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu; Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu; Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu; Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu; Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu; Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu; Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu; Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu; Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu; Thr-Glu-Glu-Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly; Thr-Glu-Glu-Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser; Thr-Glu-Glu-Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala; Thr-Glu-Glu-Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr; Glu-Glu-Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly; Glu-Glu-Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly; Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser; Glu-Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala; Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr; Gly-Leu-Gly-Phe-Leu-Tyr; Gln-Ala-Als-Lys-Glu-Met-Phe-Glu-Lys-Leu-Thr-Glu-Glu-Gly-Ser-Pro-Lys-Gly-Gln-Thr-Gly-Leu-Gly-Phe-Leu-Tyr-Ala-Ser-Gly-Leu-Gly-Val-Asn-Ser-Ser-Gln-Ala-Lys-Ala-Leu; Lys-Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly-Arg-Gly; Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly-Arg-Gly; Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly-Arg-Gly; Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly-Arg-Gly; Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly-Arg-Gly; Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly-Arg-Gly, Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly-Arg-Gly; Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly-Arg-Gly; Lys-Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly-Arg; Lys-Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly; Lys-Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly; Lys-Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His; Lys-Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu; Lys-Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His; Lys-Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu; Lys-Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln; Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly-Arg; Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly; Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly; Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His; Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu; Gln-Val-Gly-Leu-Gly-Gln-Leu-His, Gly-Leu-Gly-Gln-Leu; Val-Gly-Leu-Gly-Gln; Gln-Val-Gly-Leu-Gly-Gln; Gly-Leu-Gly-Gln-Leu-His, Gln-Val-Gly-Leu-Gly-Gln; Gly-Leu-Gly-Gln-Leu-His-Leu; Ala-Gln-Val-Gly-Leu-Gly-Gln; Glu-Asp-Leu-Ile-Gln-Tyr-Tyr-Gln-Phe-Leu-Ala-Glu-Lys-Gly-Asp-Val-Gln-Ala-Gln-Val-Gly-Leu-Gly-Gln-Leu-His-Leu-His-Gly-Gly-Arg-Gly-Val-Glu-Gln-Asn-His-Gln-Arg-Ala-Phe-Asp-Tyr-Phe; Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-

Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val, Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln-Val; Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val-Gln; Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val; Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly; Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg; Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly; Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr; Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu, Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr, Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala; Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly; Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly-Val; Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly, Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg; Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly; Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr; Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu; Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr; Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala; Gln-Ser-Gly-Leu-Gly-Met; Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala; Gly-Gln-Ser-Gly-Leu-Gly-Met; Val-Gly-Gln-Ser-Gly-Leu-Gly-Met; Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg; and Glu-Thr-Ala-Leu-His-Tyr-Phe-Lys-Lys-Ala-Ala-Asp-Met-Gly-Asn-Pro-Val-Gly-Gln-Ser-Gly-Leu-Gly-Met-Ala-Tyr-Leu-Tyr-Gly-Arg-Gly—Val-Gln-Val-Asn-Tyr-Asp-Leu-Ala-Leu-Lys-Tyr-Phe-Gln.

All of the peptides of the invention, as well as molecules substantially identical, complementary or otherwise functionally or structurally equivalent to these peptides may be used for purposes of the present invention. For example, peptides of the invention are contemplated that are phosphorylated at Tyr. In addition to full-length peptides of the invention, truncations of the peptides are contemplated in the present invention. Truncated peptides may comprise peptides of about 4 to 20, preferably 4 to 15 amino acid residues.

The truncated peptides may have an amino group (-NH2), a hydrophobic group (for example, carbobenzoxyl, dansyl, or T-butyloxycarbonyl), an acetyl group, a 9-fluorenylmethoxy-carbonyl (PMOC) group, or a macromolecule including but not limited to lipid-fatty acid conjugates, polyethylene glycol, or carbohydrates at the amino terminal end. The truncated peptides may have a carboxyl group, an amido group, a T-butyloxycarbonyl group, or a macromolecule including but not limited to lipid-fatty acid conjugates, polyethylene glycol, or carbohydrates at the carboxy terminal end.

A peptide of the invention may also include an analog of a peptide of the invention, and/or truncations thereof, which may include, but are not limited to a peptide of the invention containing one or more amino acid insertions, additions, or deletions, or both. Analogs of a peptide of the invention exhibit the activity characteristic of the peptide (e.g. interference with the interaction of a SEL-1L PTB domain binding site and a PTB domain containing protein), and may further possess additional advantageous features such as increased bioavailability, stability, or reduced host immune recognition.

One or more amino acid insertions may be introduced into a peptide of the invention. Amino acid insertions may consist of a single amino acid residue or sequential amino acids.

One or more amino acids, preferably one to ten amino acids, may be added to the right or left termini of a peptide of the invention. Deletions may consist of the removal of one or more amino acids, or discrete portions from the peptide sequence. The deleted amino acids may or may not be contiguous. The lower limit length of the resulting analog with a deletion mutation is about 7 amino acids.

It is anticipated that if amino acids are inserted or deleted in sequences outside specific interacting sequences (such as the NLGY sequence) that the resulting analog will exhibit the activity of a peptide of the invention.

The activity of a peptide of the invention may be confirmed by assaying for the ability of the peptide to interfere with the interaction of a SEL-1L PTB domain binding site and a PTB domain containing protein, or the interaction of a SEL-1L PDZ domain and a protein containing a PDZ domain binding motif.

Computer modeling techniques known in the art may also be used to observe the interaction of a peptide of the invention, and truncations and analogs thereof with a PTB domain containing protein, a SEL-1L PTB domain binding site, a SEL-1 PDZ domain, or a protein containing a PDZ domain binding motif (for example, Homology Insight II and Discovery available from BioSym/Molecular Simulations, San Diego, California, U.S.A.). If computer modeling indicates a strong interaction, the peptide can be synthesized and tested for its ability to interfere with the binding of a SEL-1L PTB domain binding site and a PTB domain containing protein, or the binding of a SEL-1 PDZ domain and a protein containing a PDZ domain binding motif, as discussed above. The peptides of the invention mediate the interactions of PTB domain containing proteins such as NUMB with SEL-1L PTB domain binding sites, and the interaction of SEL-1 PDZ domains with proteins containing a PDZ domain binding motifs such as NOTCH. The interactions regulate pathways that control gene expression, cell division, cytoskeletal architecture and cell metabolism. The peptides may therefore be used in the treatment of conditions involving perturbation of such regulatory pathways.

The invention also relates to the use of a peptide of the formula I to interfere with the interaction of a PTB domain containing protein with a SEL-1L PTB domain binding site; and, pharmaceutical compositions comprising a peptide of the formula I for inhibiting the interaction of a PTB domain containing protein with a SEL-1L PTB domain binding site. The invention also relates to the use of a peptide of the invention derived from a SEL-1L PDZ domain to interfere with the interaction of a protein containing a PDZ domain binding motif and a SEL-1L PDZ domain; and, pharmaceutical compositions comprising such a peptide for inhibiting the interaction of a protein containing a PDZ domain binding motif and a SEL-1L PDZ domain

Further, the invention relates to a method of modulating the interaction of a PTB domain containing protein with a SEL-1L PTB domain binding site comprising changing the amino acid Tyr in the consensus sequence NLGYMHE in the SEL-1L PTB domain binding site. The invention also relates to a method of modulating the interaction of a protein containing a PDZ domain binding motif and a SEL-1L PDZ domain comprising changing an amino acid in one or more of the consensus sequences GLGF, GLGQ, or GLGM of a SEL-1L PDZ domain.

The invention also includes a peptide of the invention conjugated with a selected protein, or a selectable marker to produce fusion proteins.

The peptides of the invention may be prepared using recombinant DNA methods or the peptides may be synthesized by conventional techniques as described herein. N-terminal or C-terminal fusion proteins comprising a peptide of the invention conjugated with other molecules may be prepared by fusing, through recombinant techniques, the N-terminal or C-terminal of the peptide, and the sequence of a selected protein or selectable marker with a desired biological function.

Peptides may be developed using a biological expression system. The use of these systems allows the production of large libraries of random peptide sequences and the screening of these libraries for peptide sequences that bind to particular proteins. Libraries may be produced by cloning synthetic DNA that encodes random peptide sequences into appropriate expression vectors. (see Christian et al 1992, J. Mol. Biol. 227:711; Devlin et al, 1990 Science 249:404; Cwirla et al 1990, Proc. Natl. Acad, Sci. USA, 87:6378). Libraries may also be constructed by concurrent synthesis of overlapping peptides (see U.S. Pat. No. 4,708,871).

Cyclic derivatives of the peptides of the invention are also part of the present invention. Cyclization may allow the peptide to assume a more favorable conformation for association with molecules in complexes of the invention. Cyclization may be achieved using techniques known in the art. For example, disulfide bonds may be formed between two appropriately spaced components having free sulfhydryl groups, or an amide bond may be formed between an amino group of one component and a carboxyl group of another component. Cyclization may also be achieved using an azobenzene-containing amino acid as described by Ulysse, L., et al., J. Am. Chem. Soc. 1995, 117, 8466-8467. The side chains of Tyr and Asn may be linked to form cyclic peptides. The components that form the bonds may be side chains of amino acids, non-amino acid components or a combination of the two.

It may be desirable to produce a cyclic peptide which is more flexible than the cyclic peptides containing peptide bond linkages as described above. A more flexible peptide may be prepared by introducing cysteines at the right and left position of the peptide and forming a disulphide bridge between the two cysteines. The relative flexibility of a cyclic peptide can be determined by molecular dynamics simulations.

"Peptide mimetics" are structures which serve as substitutes for peptides in interactions between molecules (See Morgan et al (1989), Ann. Reports Med. Chem. 24:243-252 for a review). Peptide mimetics include synthetic structures which may or may not contain amino acids and/or peptide bonds but retain the structural and functional features of a peptide, or enhancer or inhibitor of the invention. Peptide mimetics also include peptoids, oligopeptoids (Simon et al (1972) Proc. Natl. Acad, Sci USA 89:9367); and peptide libraries containing peptides of a designed length representing all possible sequences of amino acids corresponding to a peptide, or enhancer or inhibitor of the invention

Peptide mimetics may be designed based on information obtained by systematic replacement of L-amino acids by D-amino acids, replacement of side chains with groups having different electronic properties, and by systematic replacement of peptide bonds with amide bond replacements. Local conformational constraints can also be introduced to determine conformational requirements for activity of a candidate peptide mimetic. The mimetics may include isosteric amide bonds, or D-amino acids to stabilize or promote reverse turn conformations and to help stabilize the molecule. Cyclic amino acid

analogues may be used to constrain amino acid residues to particular conformational states. The mimetics can also include mimics of inhibitor peptide secondary structures. These structures can model the 3-dimensional orientation of amino acid residues into the known secondary conformations of proteins. Peptoids may also be used which are oligomers of N-substituted amino acids and can be used as motifs for the generation of chemically diverse libraries of novel molecules.

Peptides of the invention may be used to identify lead compounds for drug development. The structure of the peptides described herein can be readily determined by a number of methods such as NMR and X-ray crystallography. A comparison of the structures of peptides similar in sequence, but differing in the biological activities they elicit in target molecules can provide information about the structure-activity relationship of the target. Information obtained from the examination of structure-activity relationships can be used to design either modified peptides, or other small molecules or lead compounds which can be tested for predicted properties as related to the target molecule. The activity of the lead compounds can be evaluated using assays similar to those described herein.

Information about structure-activity relationships may also be obtained from co-crystallization studies. In these studies, a peptide with a desired activity is crystallized in association with a target molecule, and the X-ray structure of the complex is determined. The structure can then be compared to the structure of the target molecule in its native state, and information from such a comparison may be used to design compounds expected to possess desired activities.

The peptides of the invention may be converted into pharmaceutical salts by reacting with inorganic acids such as hydrochloric acid, sulfuric acid, hydrobromic acid, phosphoric acid, etc., or organic acids such as formic acid, acetic acid, propionic acid, glycolic acid, lactic acid, pyruvic acid, oxalic acid, succinic acid, malic acid, tartaric acid, citric acid, benzoic acid, salicylic acid, benezenesulfonic acid, and toluenesulfonic acids.

The peptides of the invention may be used to prepare monoclonal or polyclonal antibodies. Conventional methods can be used to prepare the antibodies (see Section 3 herein).

The peptides and antibodies specific for the peptides of the invention may be labeled using conventional methods with various detectable substances including enzymes, fluorescent materials, luminescent materials, biotinyl groups, and radioactive materials. Suitable enzymes, fluorescent materials, luminescent materials, and radioactive material are well known to the skilled artisan. Labeled antibodies specific for the peptides of the invention derived from the SEL-1L PTB domain binding site may be used to screen for proteins with SEL-1L PTB domain binding sites, and labeled peptides of the invention may be used to screen for PTB domain containing proteins. Labeled antibodies specific for the peptides of the invention derived from a SEL-1L PDZ domain may be used to screen for SEL-1L Related Proteins, and labeled peptides may be used to screen for proteins that contain PDZ domain binding motifs.

4.4 Compositions and Treatments

The substances or compounds identified by the methods described herein, antibodies, and antisense nucleic acid molecules of the invention, and peptides may be used for modulating the biological activity of a SEL-1L Related Protein, and they may be used in the treatment of conditions such as Alzheimer's Disease, IDDM, cancer (e.g. pancreatic cancer), neurodegenerative diseases such as Parkinson's Disease, stroke, vascular dementia, and conditions requiring modulation of fat or cholesterol metabolism such as coronary heart disease. Accordingly, the substances, antibodies, peptides, and

compounds may be formulated into pharmaceutical compositions for administration to subjects in a biologically compatible form suitable for administration *in vivo*. By "biologically compatible form suitable for administration *in vivo*" is meant a form of the active substance to be administered in which any toxic effects are outweighed by the therapeutic effects. The active substances may be administered to living organisms including humans, and animals. Administration of a therapeutically active amount of a pharmaceutical composition of the present invention is defined as an amount effective, at dosages and for periods of time necessary to achieve the desired result. For example, a therapeutically active amount of a substance may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability of antibody to elicit a desired response in the individual. Dosage regima may be adjusted to provide the optimum therapeutic response. For example, several divided doses may be administered daily or the dose may be proportionally reduced as indicated by the exigencies of the therapeutic situation.

The active substance may be administered in a convenient manner such as by injection (subcutaneous, intravenous, etc.), oral administration, inhalation, transdermal application, or rectal administration. Depending on the route of administration, the active substance may be coated in a material to protect the substance from the action of enzymes, acids and other natural conditions that may inactivate the substance.

The compositions described herein can be prepared by per se known methods for the preparation of pharmaceutically acceptable compositions which can be administered to subjects, such that an effective quantity of the active substance is combined in a mixture with a pharmaceutically acceptable vehicle. Suitable vehicles are described, for example, in Remington's Pharmaceutical Sciences (Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pa., USA 1985). On this basis, the compositions include, albeit not exclusively, solutions of the active substances in association with one or more pharmaceutically acceptable vehicles or diluents, and contained in buffered solutions with a suitable pH and iso-osmotic with the physiological fluids.

The activity of the substances, compounds, antibodies, antisense nucleic acid molecules, peptides, and compositions of the invention may be confirmed in animal experimental model systems.

4.5 Specific Applications

ALZHEIMER'S DISEASE.

The chromosomal location of *Sel-1L* makes it a candidate for an Alzheimer's Disease associated gene. *Sel-1L* is not strongly expressed in the brain which would preclude its isolation in the original screen for the chromosome XIV-linked AD gene since only a brain cDNA library was used (Sherrington et al. (Nature 375:754-760, 1995)). Since there were familial AD cases linked to chromosome XIV but with no identifiable mutations in PS1 (Sherrington et al. (Nature 375:754-760, 1995) and Alzheimer's Disease Collaborative Group (Nature Genetics 11:219-222)), it is likely that there is another gene linked to AD nearby.

Other information suggests that *Sel-1L* plays a role in AD etiology. It is similar to the yeast *hrd3* protein, identified in a screen for suppressor mutations that stabilize HMG-CoA Reductase, the key enzyme in cholesterol biosynthesis. Hrd3 probably controls the degradation of HMG-CoAR since *hrd3* mutations elevate the levels of HMG-CoAR (Hampton et al. (Mol. Biol. Cell 7:2029-44, 1996)). The link to AD is found via the *APOE* gene which has been shown to markedly affect the age of onset and the severity of AD. *ApoE*, like *hrd3*, is involved in cholesterol transport (reviewed by Hyman and Tanzi (New

Eng. J. Med. 333:1283-4, 1995)). Thus, like *apoE, Sel-1L* may have an effect on relative amounts of cholesterol in the brain which may affect the integrity of cellular membranes and the possible formation of plaques.

The effect of mutant presenilins on elevating Ab, and their localization to the golgi apparatus and ER, suggest that they may be involved in protein processing. Sel-1 also appears to be involved in degradation of notch and localizes to vesicles possibly targeted for the membrane. Hrd3 also appears to be involved in degradation of HMG-CoA Reductase. Thus, it is possible that *SEL-1L*, like PS1, might be a processing enzyme and could directly regulate Ab levels. *SEL-1L* expression in the floor plate of the neural tube and at low levels in the adult brain is consistent with a role in neural cell function.

DIABETES MELLITUS

SEL-1L maps very close to a microsatellite marker (D14s67) which has been linked to a locus affecting Insulin Dependent Diabetes Mellitus (IDDM) (Fields et al. (Genomics 33:1-8, 1996)). Thus, *SEL-1L* is also a candidate for IDDM. Other facts support this notion. *SEL-1L* is highly expressed in an insulin secreting β-islet cell line. Furthermore, linkage to both AD and IDDM does not preclude *SEL-1L*'s involvement in both disease processes since these share some common attributes. For example, both disorders manifest in a pathology of plaque formation. In AD, the plaques form in the brain and contain Ab while in IDDM, plaques containing a short peptide (beta-amylin) which is not unlike Ab form in the pancreas and this is thought to be toxic to the insulin-producing beta-islet cells (Lorenzo et al. (Nature 386:756-760)). Further similarities may be apparent when one considers that glucose metabolism is affected in IDDM and is also abnormal in AD patients (for example see Ishii et al. (J Nucl Med 38:925-928, 1997)), suggesting a common underlying cause for both disorders.

PANCREATIC CANCER

SEL-1L expression in neural and pancreatic tissues suggest a role in pancreatic cancer since a study of individuals exhibiting AD pathology in post-mortem exams also revealed the prevalence (6.7-fold) of pancreatic cancer over any other type of cancer and over the incidence of this tumor type in non-AD individuals (Burke et al. Alz. Dis. Assoc. Disor 8:22-28, 1994)). If *Sel-1L* regulates the degradation of certain key molecules (such as *notch*) involved in cellular proliferation, its malfunction might lead to uncontrolled cell growth and oncogenesis. Deregulation of notch activity has been linked to oncogenesis. In *int3* mice, the mouse mammary tumor virus is integrated into the *NOTCH 4* gene and this leads to aberrant *notch 4* expression and consequent development of mammary tumors (Robbins et al. (J. Virol. 66:2594-2599, 1992)). Furthermore, a chomosomal translocation which places the coding sequence for the cytoplasmic domain of *TAN-1* (a human *Notch* gene) downstream of the T cell receptor beta gene promoter results in the expression of a constitutively active form of notch and is associated with T cell acute lymphoblastic leukemia (T-ALL) (Ellisen et al. (Cell 66:649-661, 1991)). Furthermore, it has been recently reported that in a screen for genes that facilitate c-*myc* induced thymomas, 52% of the viral insertions used to identify these genes were in the *NOTCH 1* locus. Typically, proviral insertion led to the expression of the constitutively active cytoplasmic domain of *notch1* (Girard et al. (Genes Dev. 10:1930-1944, 1996)). Thus, it is likely that since *Sel-1L* is a regulator of notch, it also has a role in oncogenesis.

PARKINSON'S DISEASE

Sel-1L is specifically expressed in the floor plate of the neural tube at a time during which specification of the dopaminergic neurons is occurring via contact-mediated induction by the floor plate

(Hynes et al. (Cell 80:95-101, 1995)). Because dopaminergic neurons are absent in patients with Parkinson's Disease (PD), it is possible that *Sel-1L* might play a role in the etiology of PD. Further support for the model that *Sel-1L* might play a role in PD is the fact that there is a history of PD in some first degree relatives of AD patients. Thus, although AD and PD appear to have very different pathologies, there may be some common underlying factors such as the mutation of *SEL-1L* which contribute to neurodegeneration in these diseases.

CORONARY ARTERY DISEASE

The homolog of *Sel-1L* in the yeast is likely involved in the regulation of cholesterol biosynthesis (see above). *Sel-1L* may have a similar role since it is expressed in the pancreas whose function in part is to secrete various proteases and amylases and other enzymes involved in fat metabolism. The link between the manifestation of AD and the allelic variants of the *APOE* gene is also indicative of a role for fat and cholesterol metabolism in AD. In addition, post-mortem exams of patients with critical coronary artery disease revealed abundant senile plaques in the brain reminiscent of AD-like plaques (Sparks et al (Neurobiol. Aging 11:601-607, 1990)). The same factors that contribute to plaque formation in the brain may contribute to plaque formation in arteries of the cardiovascular system. This information indicates that *SEL-1L* malfunction might be a common factor in contributing to both Coronary Artery Disease and AD.

4.6 Inactivation or Alteration of Sel-1L

The invention also provides methods for studying the function of a SEL-1L Related Protein. Cells, tissues, and non-human animals lacking in *Sel-1L* expression or partially lacking in *Sel-1L* expression may be developed using recombinant expression vectors of the invention having specific deletion or insertion mutations in the *Sel-1L* gene. A recombinant expression vector may be used to inactivate or alter the endogenous gene by homologous recombination, and thereby create a *Sel-1L* deficient cell, tissue or animal.

Null alleles may be generated in cells, such as embryonic stem cells by deletion mutation. A recombinant *Sel-1L* gene may also be engineered to contain an insertion mutation which inactivates *Sel-1L*. Such a construct may then be introduced into a cell, such as an embryonic stem cell, by a technique such as transfection, electroporation, injection etc. Cells lacking an intact *Sel-1L* gene may then be identified, for example by Southern blotting, Northern Blotting or by assaying for expression of SEL-1L Protein using the methods described herein. Such cells may then be used to generate transgenic non-human animals deficient in SEL-1L Protein. Germline transmission of the mutation may be achieved, for example, by aggregating the embryonic stem cells with early stage embryos, such as 8 cell embryos, *in vitro*; transferring the resulting blastocysts into recipient females and; generating germline transmission of the resulting aggregation chimeras. Such a mutant animal may be used to define specific cell populations, developmental patterns and *in vivo* processes, normally dependent on *Sel-1L* expression.

The following non-limiting examples are illustrative of the present invention:

Example 1

In a screen of the X-Ref database supported by the National Center for Biotechnology Information (NCBI), three Expressed Sequence Tags (EST's) whose protein translation was very similar to Sel-1 were identified. In particular, one human (hEST) and two mouse (from an embryonic day 13 library mEST13, and from an embryonic day 19 library mEST19 were identified). They aligned with different regions of the Sel-1 protein (see Figure 1). mEST13 aligned in the region of Sel-1 which was similar to the

yeast protein hrd3p as well as an uncharacterized human partial cDNA in Genbank (IBD2). Because the EST's do not overlap, it is possible that they represent a family of proteins, but their expression pattern suggests that they represent one gene.

The Genbank annotations for the partial human cDNA *ibd2* included a comment that localized this gene to human chromosome XIV. PS1 maps to XIV and these two genes are genetically linked in a common pathway. The study linking *PS1* to AD also identified some AD pedigrees linked to XIV but with no detectable mutations in *PS1* [Sherrington et al. (Nature 375:754-760, 1995) and Alzheimer's Disease Collaborative Group (Nature Genetics 11:219-222)]. Furthermore, two distinct regions linked to AD were identified by microsatellite marker analysis of AD pedigrees, and *PS1* mapped to one of them (Sherrington et al. (Nature 375:754-760, 1995). ibd2 transcribed from a novel locus (designated *SEL-1L* below) may be another AD-related candidate which was missed by the screen for cDNA's in this region since only one gene was expected.

Genomic clones for *SEL-1L* were isolated and mapped via Fluorescent In Situ Hybridization (FISH) analysis to the same chromosomal interval as *PS1*-- 14q24.3-31. A more detailed mapping position for *SEL-1L* was obtained from a screen of a panel of DNA from a Radiation Hybrid (RH) map of the human genome (Stewart et al. (Gen Res 7:422-433, 1997)). This analysis placed *SEL-1L* in the general vicinity of PS1 in AD patients. In addition, *SEL-1L* also mapped very close to a microsatellite marker (D14s67) recently used to locate a locus involved in Insulin Dependent Diabetes Mellitus (IDDM) (Field et al. (Genomics 33:1-8, 1996)). These results indicate that *SEL-1L* is a candidate gene associated with IDDM and AD.

To test whether *SEL-1L* was expressed in neural tissues or in the pancreas, the site of insulin production, northern analysis of adult mouse and human tissues with all three EST's was performed. When either mEST 13 or mEST 19 were used as probes, two large widespread messages of 7.5 Kb and 4.5 Kb were identified. The two forms of the RNA were about equal in distribution. The highest expression was observed in the pancreas (at least 10-fold higher than other tissues). When hEST was used as a probe, it also exhibited widespread albeit low levels of expression, and extremely high levels in the pancreas. The predominant message detected with hEST is 7.5 Kb, like that detected with the mEST's. However, much lower levels of the 4.5 Kb message were observed with the hEST probe than with the mEST probes. This slight difference in expression between the mEST probes and hEST may indicate they represent two different genes or tissue specific splicing differences. RNA *in-situ* analysis with mEST19 supported the Northern analysis result in that mEST19 showed highest expression in the pancreas, but also some expression in the gut epithelium at E14.5 to E17.5. In addition, mEST 19 was expressed in the Dorsal Root Ganglia and at very high levels in the floor plate of the neural tube starting at E10.5 (not shown) and increasing at E11.5. The floor plate is a key inductive structure responsible for the formation of motor and dopaminergic neurons during this time (E10.5) in mouse embryonic development. In conclusion, the expression data (neural and pancreatic) supports the chromosomal mapping data in implicating *Sel-1L* in AD and IDDM.

To obtain full length cDNA sequence information, human pancreas cDNA clones which hybridized to both hEST and the mEST probes were isolated. In addition, several mouse cDNA clones from an embryonic day 16.5 library were characterized. Genomic DNA clones containing human *Sel-1L* and mouse *Sel-1L* were identified. The sequence of the mouse *Sel-1L* gene, and the deduced amino acid

sequence are shown in SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, or SEQ. ID. NO. 7, and SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, or SEQ. ID. NO.8, respectively. A sequence of the human *Sel-1L* gene, and the deduced amino acid sequence are shown in SEQ. ID. NO.9 and SEQ. ID. NO.10, respectively.

Example 2

MATERIALS AND METHODS

Cloning of mSel-1L. Three IMAGE Consortium EST clones (Genbank accession T95374, AA061453, W62650), identified by a search of the X-REF database (NCBI) using the *C. elegans* SEL-1 sequence (U50829), were obtained from Research Genetics Inc., Huntsville. These EST clones spanned different regions of the mSEL-1L cDNA. EST clones were sequenced, inserts were isolated, and used to probe a mouse E15.5 cDNA library (Clontech Inc., Palo Alto) and a human genomic library (Clontech Inc., Palo Alto). Numerous mouse cDNA clones were isolated, four of which were sequenced on both strands (ACGT Inc., Toronto, and York University, Toronto). Three human genomic clones were partially sequenced to confirm their identity (using the Fmole kit, Promega Inc., Madison) and one was used as a probe for FISH (see below).

Northern analysis and *In-situ* RNA analysis. PolyA RNA blots (Clontech Inc., Palo Alto) were hybridized with the inserts from the various EST clones, washed at 55° C at 0.2X SSC/0.1% SDS, and exposed to film using standard techniques. The probe used in Fig. 3 was derived from EST W62650 (EST clones AA061453 and T95374 gave identical results). For the *in situ* RNA analysis, frozen sections of embryos obtained from CD1 mice natural matings were prepared and hybridized as previously described (Kluppel, M., et al, 1997, Cell Growth Differ 8: 1249-56). Plug date was designated as day 0.5. The probe used was the insert from mouse EST clone AA061453.

Preparation of SEL-1L antisera and western blot analysis. A branched octamer of 15 residues corresponding to amino acids 611-625 of mSEL-1L was synthesized (University of Toronto Biotechnology Service Center), and used to immunize two New Zealand female rabbits (Reiman's Inc.) as described (Tam, J. P. Proc. Natl Acad Sci USA 85:5409-13). Rin-M, HPAF-II, ASPC-I, and COS-1 cells were obtained from ATCC and grown as recommended. The COS-1 cells were lysed in 120 mM NaCl, 50 mM Tris 7.5, 1% NP40, 25 mM NaF, and 10 mM sodium pyrophosphate with aprotinin, leupeptin, sodium vanadate and PMSF. Lysates were incubated with anti-FLAG antibody (KODAK) at 2.2 μg per immunoprecipitation and anti-mouse agarose beads (Sigma) for 1-2 hours at 4° C, washed three times in lysis buffer, and run on a 10% polyacrylamide gel. Rin-M, HPAF-II, and ASPC-I cells from near-confluent 10 cm dishes were lysed in PLC lysis buffer (Rottapel, R., et al. 1991, Mol Cell Biol 11:3043-51) and run on a 7.5% gel. After electroblotting to Protran membranes (S&S), blocking overnight in 5% non-fat dry milk (NFDM) in TBS, blots were hybridized to SEL-1L antiserum or pre-immune serum at 1:1000 in 5% NFDM. Secondary antibodies were goat anti-rabbit (Fig. 5B) or goat anti-mouse (Fig. 5A) (Biorad) used at 1:20,000 and 1:15,000, respectively, in TBS-T (TBS with 0.05% Tween 20).

Immunofluorescence. Rin-M cells were plated on gelatin-coated chamber slides (Corning) and incubated overnight. After washing in PBS, cells were fixed with 4% paraformaldehyde in PBS for 10 minutes at room temperature. Following further washing in PBS (three times), and once with 0.1% NP40 in PBS for 5 minutes, cells were then rinsed three times with PBS and incubated with SEL-1L antibody or preimmune serum at 1:100 dilution in 1% BSA and 10% normal goat serum in PBS containing 0.2% Triton X-100 for

1 hour at room temperature. Subsequently, cells were washed three times with PBS-Triton (5 minutes each) and incubated with fluorescein-conjugated goat anti-rabbit IgG (Caltag Labs) at 1:100 in BSA and normal goat serum as above. The cells were then washed three times with PBS-Triton (5 minutes each), and once with water (5 minutes). Hoechst dye (Molecular Probes) was added at 1:2000 in PBS and incubated for 1 minute followed by three washes with PBS. Slides were dried at room temperature for 30 minutes and mounted in 100% glycerol.

Expression construct and transfection into Cos-1 cells. A truncated mSEL-1L cDNA, excluding the signal sequence (amino acids 150-740), was cloned into the pFLAG-CMV2 vector (Kodak) so that the FLAG tag was in-frame at the N-terminus. Five µg of this construct was transfected into COS-1 cells using Lipofectin reagent (Gibco-BRL), according to recommended manufacturer's conditions. Cells were harvested and lysed (see above) 60 hours after transfection.

FISH analysis and Radiation Hybrid (RH) mapping. A human genomic clone for *SEL-1L* was biotinylated and hybridized to human lymphocyte DNA as previously described (Heng, H. H., Tsui, L. C. (1993) *Chromosoma* **102,** 325-32.). Amongst 100 mitotic figures, 97 of them showed signals on one pair of the chromosomes. No additional loci were detected. Superimposing Fluorescence In Situ Hybridization (FISH) signals with DAPI-stained banded chromosomes and comparing the results of 10 photos provided the detailed position (determined by SeeDNA Inc., Toronto). The medium-resolution Stanford Human Genome Center (SHGC) radiation hybrid (RH) panel was obtained from Research Genetics and used as recommended by SHGC (www.shgc.stanford.edu) using primers derived from human EST T95374 (5'-TTTAGCTTCTCAGGGAGGCC, 5'-GCAGACTTTCCTGCTGGGCA). Results were scored by positive bands on a 3% agarose gel and submitted to the SHGC RH server (www.shgc.stanford.edu/rhserver2/rhserver_form.html).

RESULTS

Cloning of a mouse homologue of the *C. elegans sel-1* gene. Because the *C. elegans* SEL-1 protein is a key regulator of NOTCH function, it was decided to clone and characterize mammalian homologues of SEL-1 (*Sel-1l* and *SEL-1L* gene symbols for the murine and human genes, respectively (for *sel-1* like, standardized by Jackson Laboratories Mouse Nomenclature Committee), and mSEL-1L, hSEL-1L, and rSEL-1L in reference to the mouse, human, and rat proteins, respectively). A comparison of the murine protein sequence, mSEL-1L, and its worm homologue SEL-1 is shown in Fig. 2A. A partial human cDNA (corresponding to residues # 589-641 in Fig. 2A) was previously called *SEL-1L/IBD2* (Biunno, I., Appierto, V., et al. (1997) *Genomics* **46,** 284-6.). The sequence of mSEL-1L is highly homologous to SEL-1 (66% similar) and exhibits similar features, including a signal peptide (as predicted by SignalP V1.1 @ www.cbs.dtu.dk/services/signal p/bg_prediction.html) immediately 3' of the putative initiator methionine (Fig. 2A). mSEL-1L also possesses a presumed PEST sequence at its C-terminus (underlined in Fig. 2A), suggesting that the levels of this protein could be tightly regulated.

The mSEL-1L cDNA (740 amino acids) encodes a protein with a predicted molecular weight of 85 kDa, the approximate size observed by western blot analysis (see below). The C-terminal region of mSEL-1L contains a domain which is highly conserved among SEL-1, hSEL-1, and mSEL-1L, and their *S. cerevisiae* homologue, Hrd3p (Fig. 2B). This region includes a glycine residue (Fig. 2B, asterisk) which is strictly conserved and required for SEL-1 function in *C. elegans* (Grant, B & Greenwald, I 1996, Genetics 143, 237-47; Grant, B & Greenwald, I, 1997 Development 124:637-44). mSEL-1L and SEL-1 are more

closely related than are mSEL-1L and Hrd3p since the latter share only 38% similarity over a 287 amino acid region (not shown).

Expression of SEL-1L in mouse and human tissues. To determine the spatial, temporal, and cell lineage pattern of expression of SEL-1L, adult mouse and human tissues were analyzed by northern blot analyses and developing mouse embryos by RNA *in-situ* analysis. SEL-1L was widely expressed in adult tissues of both mice and humans (Fig 3A & 3B), with the exception of skeletal muscle, which had low (human) to undetectable (mouse) levels of SEL-1L transcripts. Interestingly, SEL-1L was expressed at very high levels in both the adult human (Fig. 3B) and developing mouse pancreas (see below). The presence of RNA bands of sizes 7.5 and 4.5 Kb in the mouse suggests that mSEL-1L transcripts may undergo alternative splicing. Because the coding region of mSEL-1L is only 2.3 Kb (bankit192768), the mSEL-1 mRNA presumably includes long untranslated regions in the 5' and/or 3' termini.

During embryogenesis, mSEL-1L was widely expressed, with moderate expression in neural crest derived tissues, including the neural tube (arrowhead in Fig. 4A), and dorsal root ganglia (DRG) (arrow in Fig. 4A) at embryonic (E) day 11.5. *Sel-1l* expression in the floor plate of the neural tube was moderate at E10.5 (not shown), and increased to high levels by E11.5 (Fig. 4B and 4D), with a concentration of silver grains at the midline on the apical side (Fig. 4D). The notochord (derived from axial mesoderm) was negative for mSEL-1L transcripts (Fig. 4B & C). Additionally, mSEL-1L RNA in the somites was very low (not shown), suggesting that as in adult skeletal muscle (see above), mSEL-1L expression is downregulated in mesodermally-derived tissues.

As observed in adult humans, expression in the primitive mouse pancreas was very high, beginning at E14.5 (P in Fig. 4E) and persisting at E17.5 (not shown). The highest expression in the pancreas was in the acini; however, as shown below, western blot analysis suggests that SEL-1L is also expressed in islet β-cells. The epithelial cells lining the villi of the gut also expressed moderate levels of mSEL-1L transcripts at E14.5 (G in Fig. 4E) and E17.5 (not shown).

The SEL-1L protein is highly expressed in an islet tumor-derived cell line. To characterize further the mammalian SEL-1L protein, an antibody was raised in rabbits using a peptide derived from the highly conserved domain in the SEL-1L family (see Materials and Methods) that exactly matches the murine and human SEL-1L sequence. The specificity of the serum was confirmed by expressing a FLAG-tagged truncated version of mSEL-1L cDNA in COS-1 cells (amino acids 150-740). The serum recognized a specific band of the predicted size (67 kDa) in a FLAG immunoprecipitation from cells transfected with truncated mSEL-1L (Fig. 5A, lane 3), but not from cells transfected with empty vector (Fig. 5A, lane 4). Preimmune sera did not recognize any band other than immunoglobulin light chains (Fig. 5A, lanes 1 & 2). The SEL-1L sera also recognized an endogenous major band of predicted full-length size (85 KDa) in whole cell lysate from a rat islet-tumor-derived cell line (Rin-M) (Fig. 5B, lane 4). This 85 kDa band in the RinM cell lysates was not recognized by the preimmune serum (Fig. 5B, lane 1) nor in lysates of cell-lines derived from human pancreatic adenocarcinomas tested with the SEL-1L antibody (Fig. 5B, lanes 5 and 6).

Cytoplasmic localization of the SEL-1L protein. Fusion of the *C. elegans* SEL-1 signal peptide to two different reporter genes directs secretion of heterologous proteins (Grant, B & Greenwald, I, 1997 Development 124:637-44), suggesting that SEL-1 itself could be targeted for secretion. However, both a SEL-1-Green Fluorescent Protein fusion and the endogenous *C. elegans* SEL-1 protein are localized in the cytosol, possibly in intracellular vesicles (Grant, B & Greenwald, I, 1997 Development 124:637-44). To

determine the subcellular localization of the mammalian SEL-1L protein, immunofluorescence studies were performed on Rin-M cells. As shown in Fig. 5C, rat SEL-1L was localized in the cytoplasm, distributed in a puncated manner.

***SEL-1L* maps to 14q24.3-31 near an IDDM locus.** To determine whether *SEL-1L* might be implicated in any human hereditary disorders, the chromosomal location of *SEL-1L* was determined by both FISH and radiation hybrid mapping. By FISH analysis of human metaphase chromosomal spreads *SEL-1L* was mapped to the long arm of chromosome 14 in the general area of 14q24.3-31 (Fig. 6A, B, and C, and see Methods). A higher resolution physical map position was obtained by mapping *SEL-1L* on the SHGC medium-resolution RH map. *SEL-1L* was linked to D14s287 with a Lod score of 5.3 within a distance of 44.4 Centiray units, which corresponds to about 1.2 Mbp (not shown). D14s287 has been mapped on the SHGC map near D14s67 which is linked to IDDML11 (Field, L. L., et al, (1996) *Genomics* 33, 1-8) (http://www-shgc.stanford.edu/Mapping/rh/search.html).

DISCUSSION

Murine and human homologues of the *C. elegans* SEL-1 protein, a negative regulator of signaling through the Notch family members, LIN-12 and GLP-1 have been cloned. Northern and *in-situ* analyses have shown that SEL-1L was widely expressed, with particularly high levels in the floor plate of the neural tube and in developing and adult pancreas. The SEL-1L protein was present at high levels in a transformed β-islet cell line, but was not detectable in two human pancreatic adenocarcinoma cell lines. Consistent with data on SEL-1, rat SEL-1L protein was localized to the cytoplasm, possibly in secretory vesicles. Finally, the human gene was localized by FISH and RH mapping to chromosome 14q24.3-31, a region of chromosome 14 associated with predisposition to insulin dependent diabetes mellitus.

SEL-1L and the regulation of the NOTCH pathway. It has been postulated that SEL-1 might play a role in LIN-12 protein degradation (Grant, B. & Greenwald, I. (1997) *Development* 124, 637-44). This hypothesis is consistent with two observations. First, loss of function mutations in *sel-1* alleviates the phenotype of a hypomorphic *lin-12* allele, an observation consistent with SEL-1 downregulating either the level or activity of LIN-12 RNA or protein (Grant, B. & Greenwald, I. (1996) *Genetics* 143, 237-47; Grant, B. & Greenwald, I. (1997) *Development* 124, 637-44). The NOTCH proteins possess PEST sequences at their C-termini, consistent with the targeting of these proteins to the ubiquitin-mediated degradation pathway (reviewed in Weinmaster, G. (1997) *Mol Cell Neurosci* 9, 91-102). SEL-1/SEL-1L might play some role in this process. Second, the *S. cerevisiae* homologue of SEL-1, Hrd3p, was isolated in a screen for proteins that ameliorate the effects of high levels of HMG-CoA reductase. Another protein identified independently in the same screen was the regulatory subunit p97 of the 26S-proteasome particle, whose primary role is in protein degradation (Hampton, R. Y., Gardner, R. G. & Rine, J. (1996) *Mol Biol Cell* 7, 2029-44). Thus, it is possible that Hrd3p may also have a similar role as the 26S subunit in the turnover of HMG-CoA reductase. These observations imply that the related SEL-1 protein might also be involved in protein turnover.

SEL-1 and Hrd3p may exert their effects through highly specific intracellular pathways. Null mutations in either *sel-1* or *HRD3* on an otherwise wildtype background have no discernible phenotype. Furthermore, the effects of null *sel-1* and *HRD3* alleles are only observable in organisms with hypomorphic mutations in either *lin-12* and *glp-1*, or *HmgCoA reductase*, respectively (Grant, B. &

Greenwald, I. (1996) *Genetics* 143, 237-47; Grant, B. & Greenwald, I. (1997) *Development* 124, 637-44; Hampton, R. Y., Gardner, R. G. & Rine, J. (1996) *Mol Biol Cell* 7, 2029-44).

SEL-1/SEL-1L may exert their negative effects on Notch signaling through the processing and/or downstream signaling of NOTCH proteins. Cleavage of NOTCH by the metalloprotease Kuzbanian (KUZ) in golgi-associated vesicles is thought to be critical for activity or stabilization of NOTCH (Pan, D. & Rubin, G. M. (1997) *Cell* 90, 271-80). As SEL-1/SEL-1L appear to localize to intracellular vesicles, these proteins might participate in NOTCH processing. In the worm, the genetic evidence suggests that *sel-1* and *lin-12* act within the same cell. Analysis of *Sel-1L* expression in the mouse indicates that while the murine *Notch-1* and *Sel-1L* genes show extensive overlap in expression, *Sel-1L* is also expressed in tissues not known to express Notch. For example, moderate levels of mSEL-1L transcripts coincide with those of NOTCH-1 in the DRG and neural tube at E11.5, and hSEL-1L coincides with NOTCH-1 in adult pancreas (Ellisen, L. Wet al. (1991) *Cell* 66, 649-61; Reaume, A. G. et al. (1992) *Dev Biol* 154, 377-8722; Del Amo, F. et al. (1992) *Development* 115, 737-44). However, while mSEL-1L transcripts are highly expressed in the floor plate at E10.5-E11.5, NOTCH transcripts have not been detected in this tissue (Williams, R., et al. (1995) *Mech Dev* **53**, 357-68, Lardelli, M. et al. (1994) *Mech Dev* 46, 123-36). Therefore, SEL-1L may not be restricted in its mode of action to the Notch pathway.

High mSEL-1L RNA correlates with secretory organs. The highest levels of mSEL-1L RNA were observed in pancreas, gut, and floor plate, all organs which are known to secrete various protein factors. The pancreas secretes exocrine and endocrine factors, while the gut secretes digestive enzymes (Slack, J. M. (1995) *Development* 121, 1569-80). The floor plate produces diffusible factors such as Sonic Hedgehog (SHH) and the Netrins, which mediate both long- and short-range induction and guidance of neurons in the central nervous system (Kennedy, T. E., et al. (1994) *Cell* 78, 425-35; Marti, E. et al. (1995) *Development* 121, 2537-47). The clustering of silver grains corresponding to mSEL-1L transcripts over the apical side of the floor plate and the lumen of the neural tube is consistent with their localization to budding vesicles from the floor plate. This spatial and temporal localization of mSEL-1L transcripts coincides with that observed for SHH RNA (Marti, E et al. (1995) *Development* 121, 2537-47).

The SEL-1L protein is absent from human pancreatic adenocarcinomas. High expression of mSEL-1L and rSEL-1L was observed in acini of the developing mouse pancreas and in a rat β-islet tumor-derived cell line, respectively, suggesting that SEL-1L is expressed in both the exocrine and endocrine cells of the pancreas. The pancreas contains a third cell type involved in forming the ducts into which the exocrine factors secreted by the acini are collected. The ducts and acini can give rise to pancreatic adenocarcinomas (Parsa, I., et al. (1985) *Cancer Res* 45, 1285-90). hSEL-1L protein was not detected in two established pancreatic adenocarcinoma tumor-derived human cell-lines (HPAF-II and ASPC-1). Thus, it is possible that hSEL-1L expression is downregulated in a fraction of these tumors. Activation of the Notch signaling pathway has been directly implicated in neoplastic transformations (reviewed in (Gridley, T. (1997) *Mol Cell Neurosci* 9, 103-8). Thus, the absence of hSEL-1L, a human homologue of a known negative regulator of NOTCH signaling, in pancreatic neoplasms raises the possibility that the Notch pathway is involved in pancreatic adenocarcinomas.

***SEL-1L* maps close to the diabetes locus IDDML11.** The human *SEL-1L* gene was localized to chromosome 14q24.3-31, in close proximity to IDDML11. IDDM is an autoimmune disease which culminates in immune system destruction of pancreatic beta-islet cells (Boitard, C., et al. (1997) *Horm Res*

48 Suppl 4, 58-63). It is unclear whether IDDM is caused by the presentation of an aberrant protein on β-islet cell surfaces which marks these cells for destruction by infiltrating lymphocytes, or the misrecognition of normal islet antigens by abnormal T-cells. The high level of SEL-1L expression in the pancreas, and its putative role in protein processing/degradation, indicates that SEL-1L may be involved in targeting the beta-islet cells for destruction by the immune system.

Example 3

PTB binding site

Peptide binding studies have shown that a conserved region of SEL-1L (amino acids #641-650 in the mouse sequence) contains a PTB binding site. This corresponds to the sequence NLGY. In-vitro binding studies using a Bia-core machine and fluorescence depolarization techniques have demonstrated that NUMB PTB domain can bind the SEL-1L peptide in a specific manner with a Kd of 7.6 μM (Figure 8). The NUMB PTB domain (both mammalian and Drosophila isoforms) can also pull-down normal endogenous SEL-1L protein from extracts of an islet tumor cell line. Finally, preliminary experiments suggest that the two proteins can associate together when co-expressed in human 293 cells. NUMB, like SEL-1L, is a negative regulator of the NOTCH pathway. This association indicates that these two regulators might cooperate in modulating NOTCH activity. The islet tumor cell line has been found to make endogenous NUMB, SEL-1L, and NOTCH.

PDZ domains

Based on secondary structure predictions and sequence comparisons it has been determined that SEL-1L contains three putative PDZ domains containing the consensus sequences: GLGF at position #231; GLGQ at position #349; GLGM at position #423. The NOTCH receptor contains a PDZ binding motif at its C-terminus which has been shown to associate with Dishevelled. SEL-1L may interact with the NOTCH receptor via the PDZ domains and corresponding binding site resulting in regulatory consequences.

Having illustrated and described the principles of the invention in a preferred embodiment, it should be appreciated to those skilled in the art that the invention can be modified in arrangement and detail without departure from such principles. All modifications coming within the scope of the following claims are claimed.

All publications, patents and patent applications referred to herein are incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety.

The application contains sequence listings which form part of the application.

Detailed Figure Legends:

Figure 2. Sequence alignment of SEL-1L family members. Figure 2A. mSEL-1L (bankit192768), top line, is compared to SEL-1 (Genbank accession U50829), bottom line. The numbering of amino acids is shown only for mSEL-1L. The putative signal sequence of mSEL-1L is designated with a line over the sequence. Mutation of the glycine residue (marked above with a star) to either an Arginine or Glutamic acid inactivates *sel-1* in *C. elegans*. A putative PEST sequence in mSEL-1L is underlined. In two out of three mSEL-1L cDNAs, the open reading frame (ORF) terminated at codon 740 as shown. A third cDNA clone encoded an ORF that terminated at residue 627, excluding a short stretch of conserved residues and the putative PEST sequence (bankit192773). Figure 2B. A comparison of a C-terminal conserved region among SEL-1L family members. The region shown corresponds to residues 584-642 in mSEL-1L. The hSEL-1L sequence corresponds to a partial human cDNA (Genbank accession U11037) (18). The hrd3p protein was previously reported (PIR database S48558). The conserved glycine residue mutated in *C. elegans* is marked with a star.

Figure 3. hSEL-1L and mSEL-1L transcripts in adult tissues. FIGURE 3A. Top panel. Mouse tissue poly-A RNA blot was hybridized with a mSEL-1L probe. h, heart; b, brain; s, spleen; lu, lung; li, liver; sk, skeletal muscle; k, kidney, t, testis. Two RNA species, of sizes 7.5 Kb and 4.5 Kb were observed. Bottom panel. The same blot was hybridized with glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as a control for RNA loading. Figure 3B. Top panel. Human tissue poly-A RNA blot was hybridized with the same mSEL-1L probe used in A. h, heart; b, brain; pl, placenta; lu, lung; li, liver; sk, skeletal muscle; k, kidney; p, pancreas. Only the 7.5 Kb band was detected in human tissues. Bottom panel. The same blot hybridized with GAPDH.

Figure 4. mSEL-1L expression in developing mouse embryos. RNA *in situ* analysis was performed using an EST derived mSEL-1L probe on frozen sections of mouse embryos. Figure 4A. Dark-field image (10X magnification) of a cross-section through an E11.5 mouse embryo showing moderate levels of mSEL-1L expression in the DRG's (white arrow) and low levels throughout the neural tube (arrowhead). Figure 3B. Dark-field image (10X magnification) of a cross-section through an E11.5 embryo showing high expression in the floor plate of the neural tube (the bright spot in the middle of the section) and lack of expression in the notochord (see Figure 4C). Figure 4C. Bright-field image (10X magnification) of the same section shown in panel B. fp, floor plate. The black arrow indicates the notochord, which in B is negative for mSEL-1L transcripts. Figure 4D. Bright field closeup (40X magnification) of the same section shown in B and C, showing the silver grains condensation over the midline on the apical side of the floor plate. Figure 4E. Dark-field (10X magnification) of a saggital section through an E14.5 embryo showing moderate expression of mSEL-1L in the epithelial layer of the villi of the gut and high expression in the pancreas. G, gut; P, pancreas. The sense probe hybridized to adjacent sections showed no staining.

Figure 5. Expression and subcellular localization of the SEL-1L protein. Figure 5A. Western blot confirming the SEL-1L sera. Lysates of Cos-1 cells transfected with FLAG-tagged truncated mSEL-1L cDNA (residues 150-740) (lanes 1 and 3) or empty vector (lanes 2 and 4) were immunoprecipiated with FLAG antibody and the western was hybridized with immune SEL-1L sera (lanes 3 and 4) or preimmune sera (lanes 1 and 2). The predicted size of the tagged truncated mSEL-1L cDNA is approximately 67 kDa, the observed size recognized by the serum in the cDNA transfected lane (lane 3). The lower molecular weight bands are probably the immunoglobulin light chains. Molecular weight standards are indicated on

the right. Figure 5B. Western blot showing endogenous SEL-1L in rat islet cells but not in human pancreatic adenocarcinoma cell lines. Whole cell lysates from RinM (lanes 1 and 4), HPAFII (lanes 2 and 5), and ASPC1 (lanes 3 and 6) cells were blotted with SEL-1L serum (lanes 4-6) or preimmune serum (lanes 1-3). A band of predicted size (85 kDa) was present in the RinM sample (lane 4) but was not recognized by the preimmune sera (lane 1) nor by the immune serum in the pancreatic adenocarcinoma cell lines (lanes 5 and 6). The high and low molecular weight bands present in all the lanes are nonspecific. Molecular weight standards are indicated on the right. Figure 5C. Immunofluorescence with RinM cells showing the cytoplasmic location of the SEL-1L protein. RinM cells incubated with preimmune serum or secondary antibody alone showed no staining (not shown).

Figure 6. Genomic localization of *SEL-1L* to chromosome 14q24.3-31 by FISH. Figure 6A. Fluorescence staining shows hybridization with a genomic *SEL-1L* probe to human metaphase chromosomes (white arrows). Figure 6B. Dapi-stained chromosomes show that the signal in A corresponds to the long arm of chromosome 14. Figure 6C. A giemsa-band diagram of human chromosome 14 showing the clustering of the fluorescence signal in A to an interval between 14q24.3 and 31.

WE CLAIM:

1. An isolated nucleic acid molecule which comprises:

i) a nucleic acid sequence encoding a protein having substantial sequence identity preferably at least 70% sequence identity, with the amino acid sequence of SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10;

ii) a nucleic acid sequence encoding a protein comprising the amino acid sequence of SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10;

iii) nucleic acid sequences complementary to (i);

iv) a degenerate form of a nucleic acid sequence of (i);

v) a nucleic acid sequence capable of hybridizing under stringent conditions to a nucleic acid sequence in (i), (ii) or (iii);

vi) a nucleic acid sequence encoding a truncation, an analog, an allelic or species variation of a protein comprising the amino acid sequence of SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10; or

vii) a fragment, or allelic or species variation of (i), (ii) or (iii).

2. An isolated nucleic acid molecule as claimed in claim 1 which comprises:

i) a nucleic acid sequence comprising the sequence of SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ. ID. NO. 7, or SEQ. ID. NO. 9 wherein T can also be U;

ii) nucleic acid sequences complementary to (i), preferably complementary to the full nucleic acid sequence of SEQ. ID. NO. 1, SEQ. ID. NO. 3, SEQ. ID. NO. 5, SEQ. ID. NO. 7, or SEQ. ID. NO. 9;

iii) a nucleic acid capable of hybridizing under stringent conditions to a nucleic acid of (i) or (ii) and preferably having at least 18 nucleotides; or

iv) a nucleic acid molecule differing from any of the nucleic acids of (i) to (iii) in codon sequences due to the degeneracy of the genetic code.

3. A recombinant expression vector adapted for transformation of a host cell comprising a nucleic acid molecule as claimed in claim 1 or 2 and one or more transcription and translation elements linked to the nucleic acid molecule.

4. A transformed host cell comprising a nucleic acid as claimed in claim 1or 2.

5. Transgenic non-human mammals whose germ cells and somatic cells contain a recombinant molecule comprising a nucleic acid molecule as claimed in claim 1 or 2.

6. A method for preparing a SEL-1L protein comprising (a) transferring a recombinant expression vector as claimed in claim 3 into a host cell; (b) selecting transformed host cells from untransformed host cells; (c) culturing a selected transformed host cell under conditions which allow expression of a SEL-1L protein; and (d) isolating the SEL-1L protein.

7. An isolated SEL-1L protein comprising the amino acid sequence as shown in SEQ. ID. NO. 2, SEQ. ID. NO. 4, SEQ. ID. NO. 6, SEQ. ID. NO. 8, or SEQ. ID. NO. 10.

8. Antibodies having specificity against an epitope of a SEL-1L protein as claimed in claim 7.

9. A probe comprising a nucleic acid molecule as claimed in claim 1 or 2.

10. A method for identifying a substance which binds to a protein as claimed in claim 7 comprising (a) reacting the protein with at least one substance which potentially can bind with the protein, under conditions which permit the binding of the substance and protein; and (b) assaying for binding; wherein the detection of binding indicates that the test substance binds to the protein.

11. A method for evaluating a compound for its ability to modulate the biological activity of a SEL-1L protein as claimed in claim 7 which comprises (a) reacting the protein, with a substance which binds to the protein and a test compound under conditions which permit binding of the substance and protein, and (b) detecting binding, wherein the detection of increased or decreased binding relative to binding in the absence of the test compound indicates that the test compound modulates the biological activity of the SEL-1L protein.

12. A method for screening for a substance that modulates the interaction of a SEl-1L PTB domain binding site and a PTB domain containing protein comprising

a. reacting a SEL-1L PTB domain binding site, a PTB domain containing protein, and a test substance, under conditions which permit the binding of the SEL-1L PTB domain binding site and the PTB domain containing protein; and

b. assaying for inhibition or stimulation of the binding of the SEL-1L PTB domain binding site and the PTB domain containing protein

wherein detection of inhibition or stimulation of binding relative to binding in the absence of the test substance indicates that the substance modulates the interaction of a SEL-1L PTB domain binding site and a PTB domain containing protein.

13. A method for screening for a substance that modulates the interaction of a SEL-1L PDZ domain and a protein containing a PDZ domain binding motif comprising

(a) reacting a SEL-1L PDZ domain, a protein containing a PDZ domain binding motif, and a test substance, under conditions which permit the binding of the SEL-1L PDZ domain and the protein containing a PDZ domain binding motif; and

(b) assaying for inhibition or stimulation of the binding of the SEL-1L PDZ domain and the protein containing a PDZ domain binding motif;

wherein detection of inhibition or stimulation of binding relative to binding in the absence of the test substance indicates that the substance modulates the interaction a SEL-1L PDZ domain and a protein containing a PDZ domain binding motif.

14. A peptide which interferes with the interaction of a Sel-1L PTB domain binding site with a PTB domain containing protein.

15. A peptide as claimed in claim 14 which comprises Asn-Leu-Gly-Tyr-Met-His-Glu.

16. A peptide as claimed in claim 14 of the formula I

$$X_n\text{- Asn-Leu-Gly-Tyr-Met-His-Glu -}Y_n \qquad \text{I}$$

wherein X and Y represent any amino acid and n is 0 to 10.

17. A peptide which interferes with the interaction of a Sel-1L PDZ domain with a protein containing a PDZ domain binding motif.

18. A peptide as claimed in claim 17 which comprises one or more of the amino acid sequences GLGF, GLGQ, or GLGM.

19. A composition comprising a protein as claimed in claim 7 and a pharmaceutically acceptable carrier, excipient or diluent.

20. A composition comprising a peptide as claimed in any one of claims 14 to 18 and a pharmaceutically acceptable carrier, excipient or diluent.

21. A composition comprising a compound or substance identified using a method as claimed in any one of claims 10 to 13, and a pharmaceutically acceptable carrier, excipient or diluent.

22. A method for treating or preventing Alzheimer's Disease, diabetes, cancer, stroke, vascular dementia, Parkinson's Disease, or coronary heart disease comprising administering to a patient in need thereof, a composition as claimed in claim 19, 20, or 21.

Figure 1

```
                1                                                    50
Yscllhom.Frg    .......... .......... .......... .......... ..........
   HIP-1.Frg    .......... .......... .......... .......... ..........
Mest-E13.Frg    .......... .......... .......... .......... ..........
 Humibd2.Frg    .......... .......... .......... .......... ..........
   Sel-1.Frg    ..IQRLRLTD ETDPTIHMQP GSAPLESNLL EYYKMLADKG DTSAQLGLGQ
Mest-E19.Frg    .VVQRIRLPD EVENPGMNS. ..GMLEEDLI QYYQFLAEKG DVQAQVGLGQ

               51                                                   100
Yscllhom.Frg    .......... .......... .......... .......... ..........
   HIP-1.Frg    .......... .......... .......... .......... ..........
Mest-E13.Frg    .......... .......... .......... .......... ..........
 Humibd2.Frg    .......... .......... .......... .......... ..........
   Sel-1.Frg    IYLAGGRGLN QNFELAFRYL LAAAESGSAD ALTYLGKMYL DGTPFTPKDY
Mest-E19.Frg    LHLHGGRGVE QNHQRAFDYF NLAANAGN.. .......... ..........

              101                                                   150
Yscllhom.Frg    .......... .......... .......... .......... ..........
   HIP-1.Frg    .......... .......... .......... .......... ..........
Mest-E13.Frg    .......... .......... .......... .......... ..........
 Humibd2.Frg    .......... .......... .......... .......... ..........
   Sel-1.Frg    QKSFEYLMKS ADKSSPSAQA VLGAMYMKGK GVKKNYEKAL KLLTLSADKK
Mest-E19.Frg    .......... .......... .......... .......... ..........

              151                                                   200
Yscllhom.Frg    .......... .......... .......... .......... ..........
   HIP-1.Frg    ...ILSLLSR VIFFFLTDGI GVKRDYKQAL KYFNLASQGG HILAFYNLAQ
Mest-E13.Frg    .......... .......... .......... .......... ..........
 Humibd2.Frg    .......... .......... .......... .......... ..........
   Sel-1.Frg    NADGQMYLAE LHYKGVPTNK GVHRDFKKSV KLYQLASQNG HILAYYNLAQ
Mest-E19.Frg    .......... .......... .......... .......... ..........

              201                                                   250
Yscllhom.Frg    .......... .......... .......... .......... ..........
   HIP-1.Frg    MHASGRGVMR SCHTAVEV.R SFLPACV... .......... ..........
Mest-E13.Frg    .......... .......... .......... .......... ..........
 Humibd2.Frg    .......... .......... .......... .......... ..........
   Sel-1.Frg    MHAAGTGVPR SCSHAVDLFK SVAERGKWGE RLMEAHSAYK DNRVDEAAMK
Mest-E19.Frg    .......... .......... .......... .......... ..........

              251                                                   300
Yscllhom.Frg    .......... .......... .......... .......... ..........
   HIP-1.Frg    .......... .......... .......... .......... ..........
Mest-E13.Frg    .......... .......... .......... .........RH RLHWNRAASQ
 Humibd2.Frg    .......... .......... .......... .......... ..........
   Sel-1.Frg    YLFMAELGYE VAQTNLAYIL DRGEATSLFS GPKDNNMERA FLNWQRSANQ
Mest-E19.Frg    .......... .......... .......... .......... ..........

              301                                                   350
Yscllhom.Frg    ......VVAG DIYFQMQ... .NYSKAMALY QGAALKY.SI QAIWNLGYMH
   HIP-1.Frg    .......... .......... .......... .......... ..........
Mest-E13.Frg    GYTVARIKLG DYHFYGFGSD VDYETAFIHY RLASEQQHSA QAMFNLGYMH
 Humibd2.Frg    .......... .SISMGLDTD VDYETAFIHY RLASEQQHSA QAMFNLGYMH
   Sel-1.Frg    EYAAARVKLG DYYYYGLGTE VDHSLAFSNY KMAVDRHGVA QAMFNLGYMH
Mest-E19.Frg    .......... .......... .......... .......... ..........

              351                                                   400
Yscllhom.Frg    EHGLGVNRDF HLAKRYYDQV SEHDHRFYLA SKLSVLKLHL KSWLT.....
   Hip-1.Frg    .......... .......... .......... .......... ..........
Mest-E13.Frg    EKGLGIKQDI HLAKRFYDMA AEASPDAQVP VFLALCKLGV VYFLQYIREA
 Humibd2.Frg    EKGLGIKQDI HLAKRFYDMA A.VSQPRCTS SSLPSPLQIG HRLFLAVHTG
   Sel-1.Frg    EVGEGITRDL YLAKRFYDQA IEHSQDAYMP SKLALAKLAF VFYLEELNKL
Mest-E19.Frg    .......... .......... .......... .......... ..........

              401                                  438
Yscllhom.Frg    .......... .......... .......... ........
   HIP-1.Frg    .......... .......... .......... ........
Mest-E13.Frg    N......... .......... .......... ........
 Humibd2.Frg    NKHSRYVLPT *......... .......... ........
   Sel-1.Frg    PLISFMEKTV GPRWDAILMT VSALVPLFLF WRHRQNDN
Mest-E19.Frg    .......... .......... .......... ........
```

Figure 2A

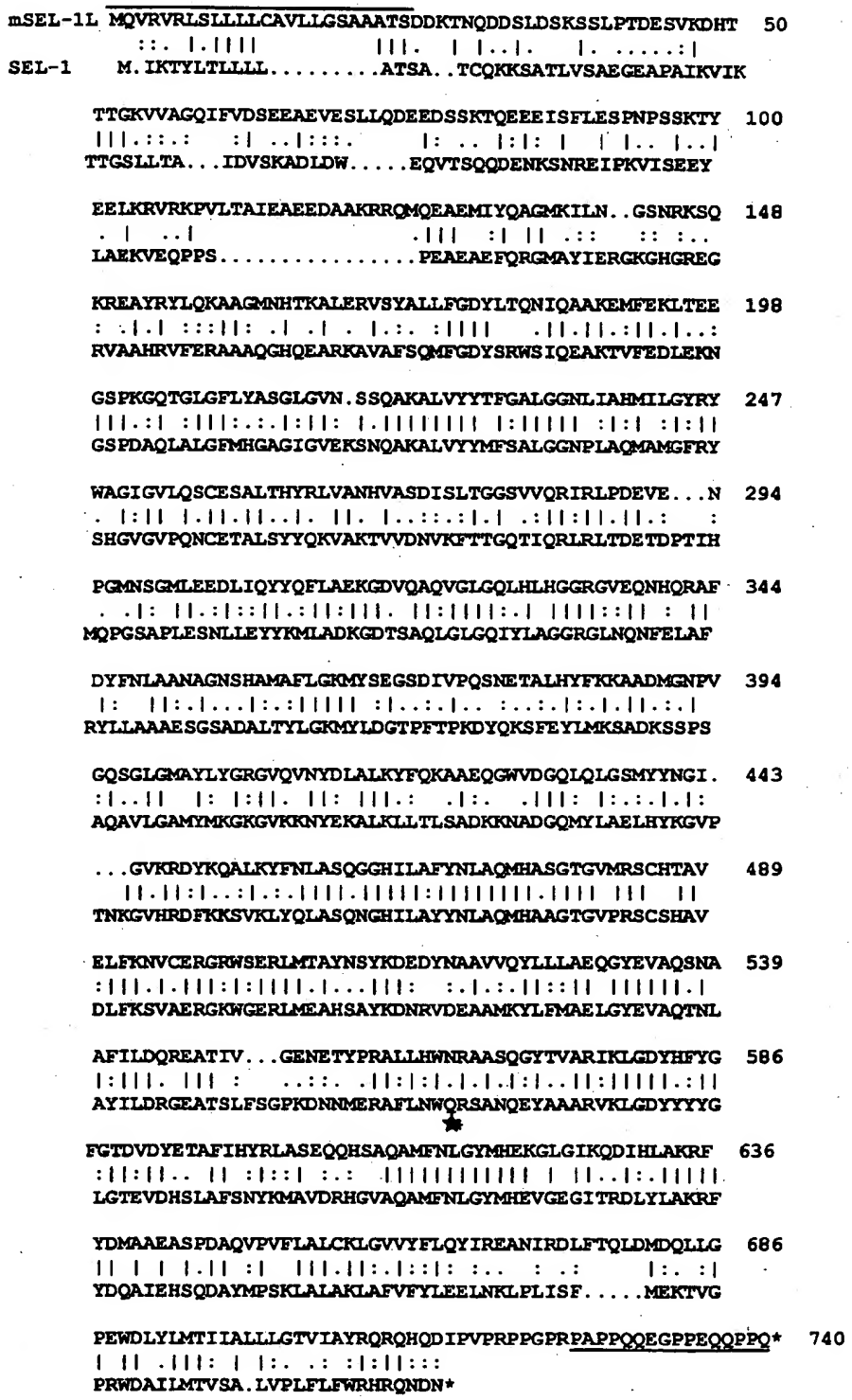

mSEL-1L MQVRVRLSLLLLCAVLLGSAAATSDDKTNQDDSLDSKSSLPTDESVKDHT 50
SEL-1 M.IKTYLTLLLL.....ATSA..TCQKKSATLVSAEGEAPAIKVIK
TTGKVVAGQIFVDSEEAEVESLLQDEEDSSKTQEEEISFLESPNPSSKTY 100
TTGSLLTA...IDVSKADLDW.....EQVTSQQDENKSNREIPKVISEEY
EELKRVRKPVLTAIEAEEDAAKRRQMQEAEMIYQAGMKILN..GSNRKSQ 148
LAEKVEQPPS.....PEAEAEFQRGMAYIERGKGHGREG
KREAYRYLQKAAGMNHTKALERVSYALLFGDYLTQNIQAAKEMFEKLTEE 198
RVAAHRVFERAAAQGHQEARKAVAFSQMFGDYSRWSIQEAKTVFEDLEKN
GSPKGQTGLGFLYASGLGVN.SSQAKALVYYTFGALGGNLIAHMILGYRY 247
GSPDAQLALGFMHGAGIGVEKSNQAKALVYYMFSALGGNPLAQMAMGFRY
WAGIGVLQSCESALTHYRLVANHVASDISLTGGSVVQRIRLPDEVE...N 294
SHGVGVPQNCETALSYYQKVAKTVVDNVKFTTGQTIQRLRLTDETDPTIH
PGMNSGMLEEDLIQYYQFLAEKGDVQAQVGLGQLHLHGGRGVEQNHQRAF 344
MQPGSAPLESNLLEYYKMLADKGDTSAQLGLGQIYLAGGRGLNQNFELAF
DYFNLAANAGNSHAMAFLGKMYSEGSDIVPQSNETALHYFKKAADMGNPV 394
RYLLAAAESGSADALTYLGKMYLDGTPFTPKDYQKSFEYLMKSADKSSPS
GQSGLGMAYLYGRGVQVNYDLALKYFQKAAEQGWVDGQLQLGSMYYNGI. 443
AQAVLGAMYMKGKGVKKNYEKALKLLTLSADKKNADGQMYLAELHYKGVP
...GVKRDYKQALKYFNLASQGGHILAFYNLAQMHASGTGVMRSCHTAV 489
TNKGVHRDFKKSVKLYQLASQNGHILAYYNLAQMHAAGTGVPRSCSHAV
ELFKNVCERGRWSERLMTAYNSYKDEDYNAAVVQYLLLAEQGYEVAQSNA 539
DLFKSVAERGKWGERLMEAHSAYKDNRVDEAAMKYLFMAELGYEVAQTNL
AFILDQREATIV...GENETYPRALLHWNRAASQGYTVARIKLGDYHFYG 586
AYILDRGEATSLFSGPKDNNMERAFLNWQRSANQEYAAARVKLGDYYYYG
FGTDVDYETAFIHYRLASEQQHSAQAMFNLGYMHEKGLGIKQDIHLAKRF 636
LGTEVDHSLAFSNYKMAVDRHGVAQAMFNLGYMHEVGEGITRDLYLAKRF
YDMAAEASPDAQVPVFLALCKLGVVYFLQYIREANIRDLFTQLDMDQLLG 686
YDQAIEHSQDAYMPSKLALAKLAFVFYLEELNKLPLISF.....MEKTVG
PEWDLYLMTIIALLLGTVIAYRQRQHQDIPVPRPPGPRPAPPQQEGPPEQQPPQ* 740
PRWDAILMTVSA.LVPLFLFWRHRQNDN*

Figure 2B

```
                                                        *
mSEL-1L  FYGFGSDVDYETAFIHYRLASEQQHSAQAMFNLGYMHEKGLGIKQDIHLAKRFYDMAAE
hSEL-1L  SMGLDTDVDYETAFIHYRLASEQQHSAQAMFNLGYMHEKGLGIKQDIHLAKRFYDMAAV
SEL-1    YYGLGTEVDHSLAFSNYKMAVDRHGVAQAMFNLGYMHEVGEGITRDLYLAKRFYDQAIE
Hrd3p    YFQMQNYSKAM-ALYQG-AALKYS--IQAIWNLGYMHEHGLGVNRDGHLAKRYYDQVSE
```

Figure 3A

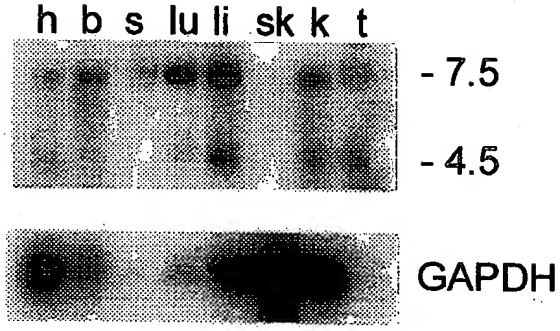


Figure 3B

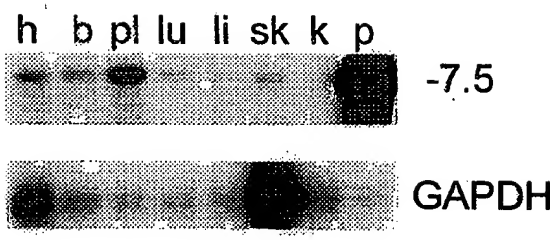


Figure 4A

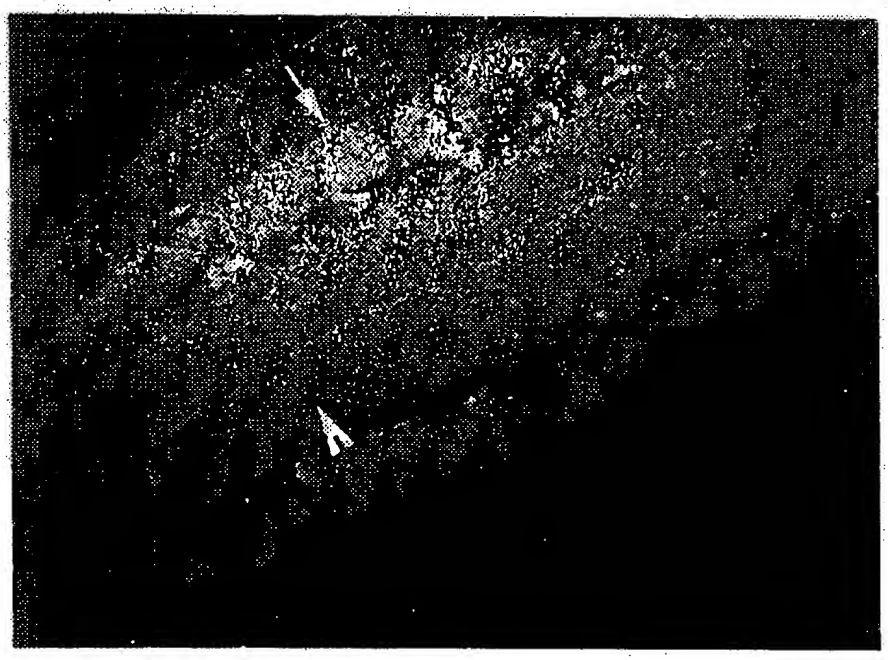


Figure 4B

Figure 4C

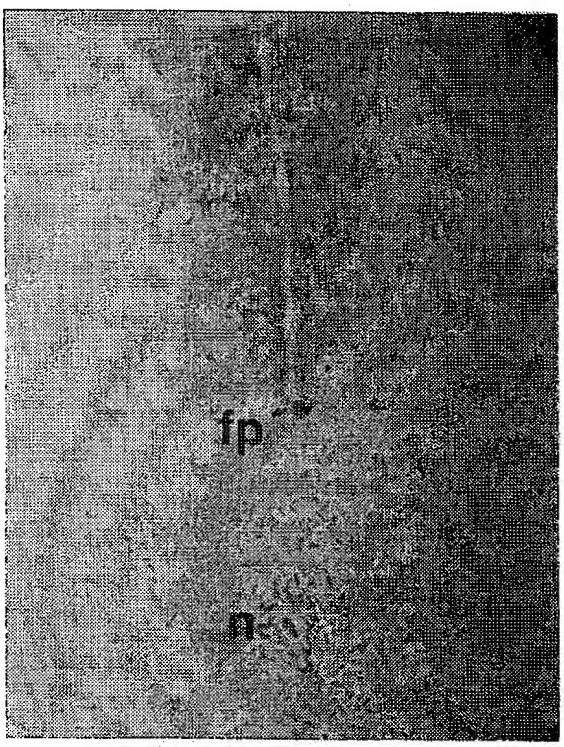


Figure 4D

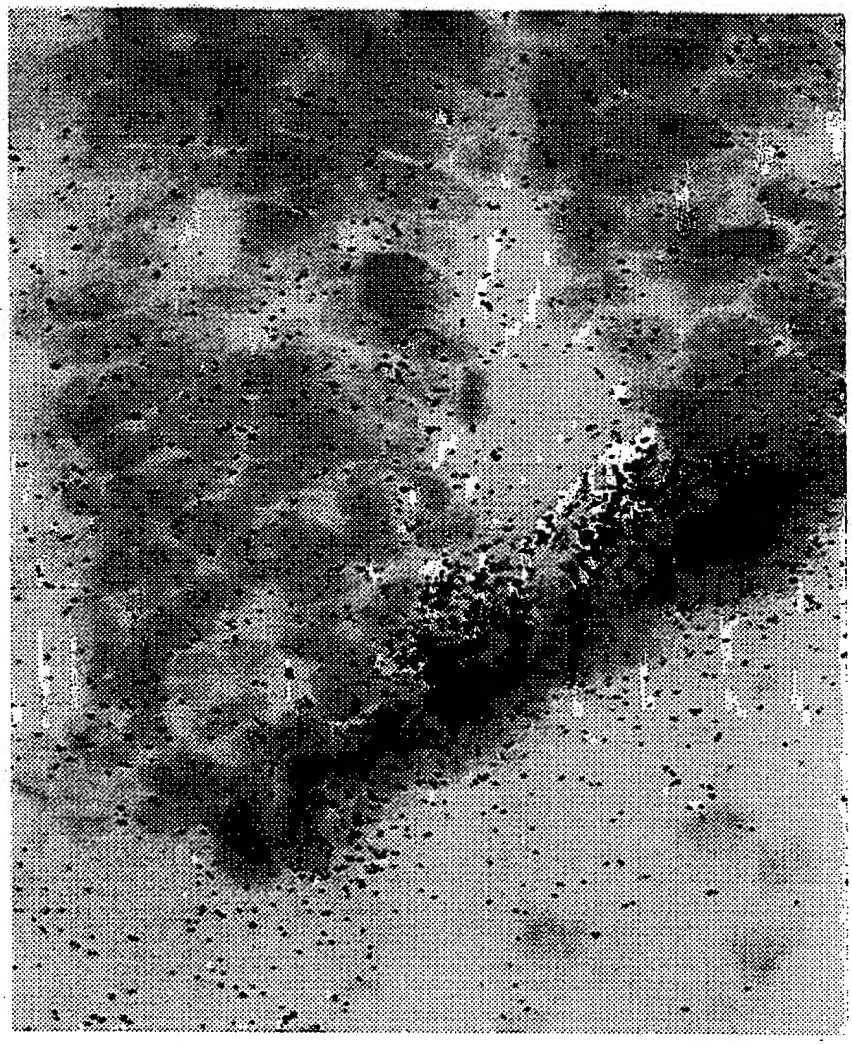


Figure 4E

Figure 5A

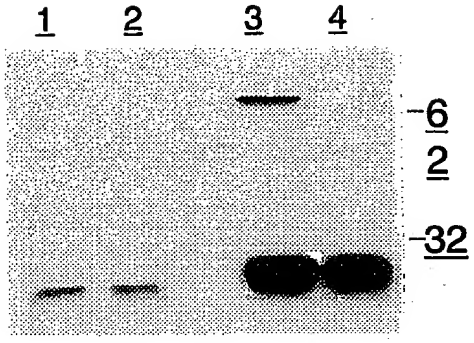


Figure 5B

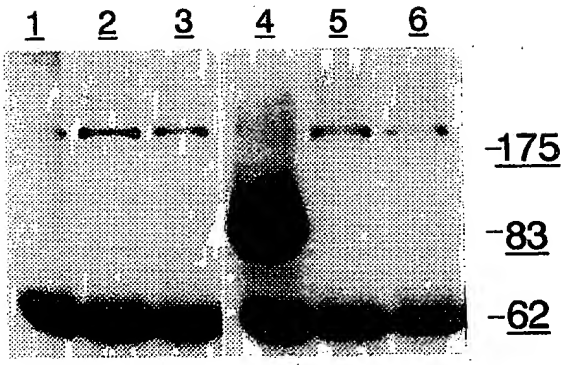


Figure 5C

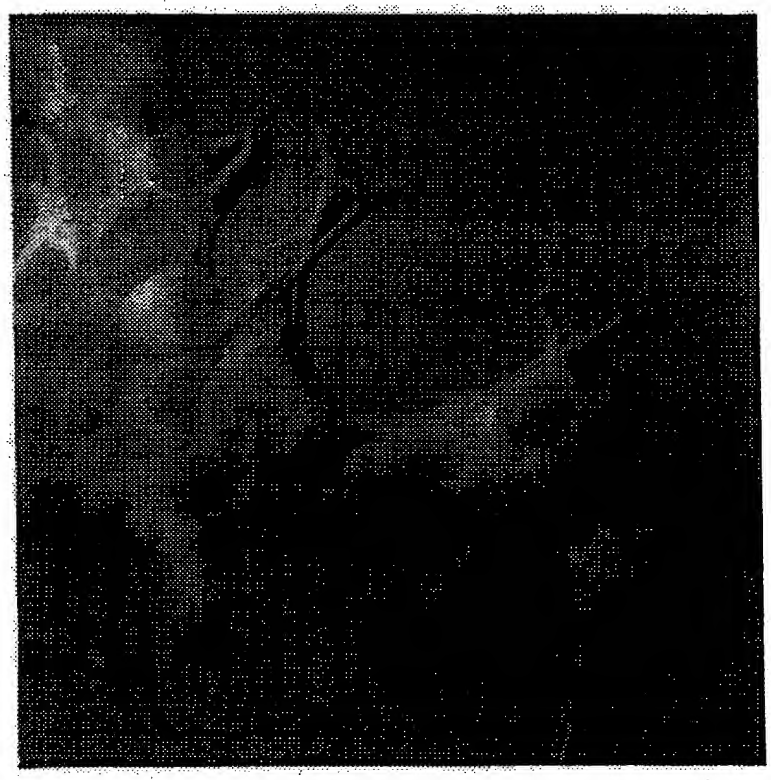


Figure 6A

Figure 6B

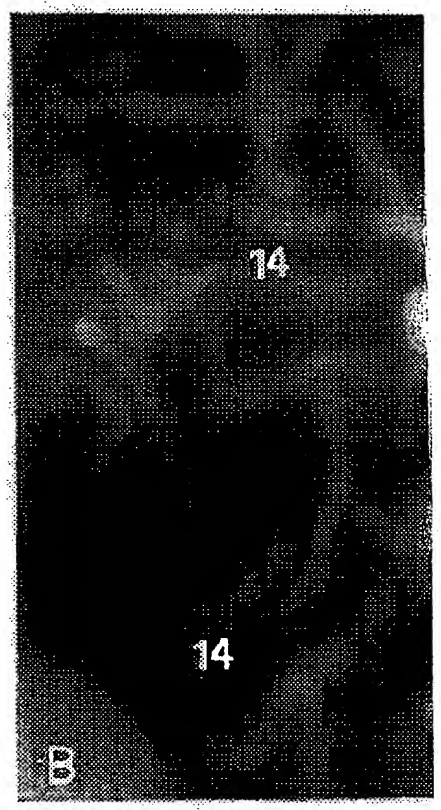


Figure 6C

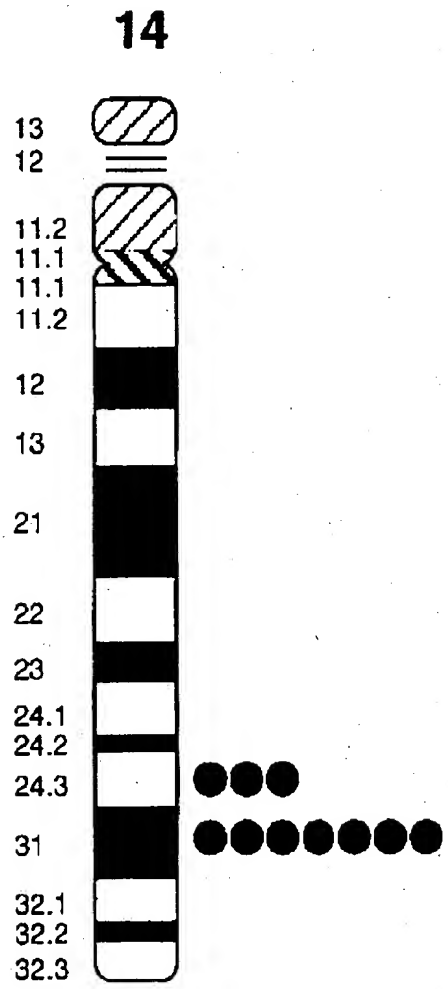


Figure 7

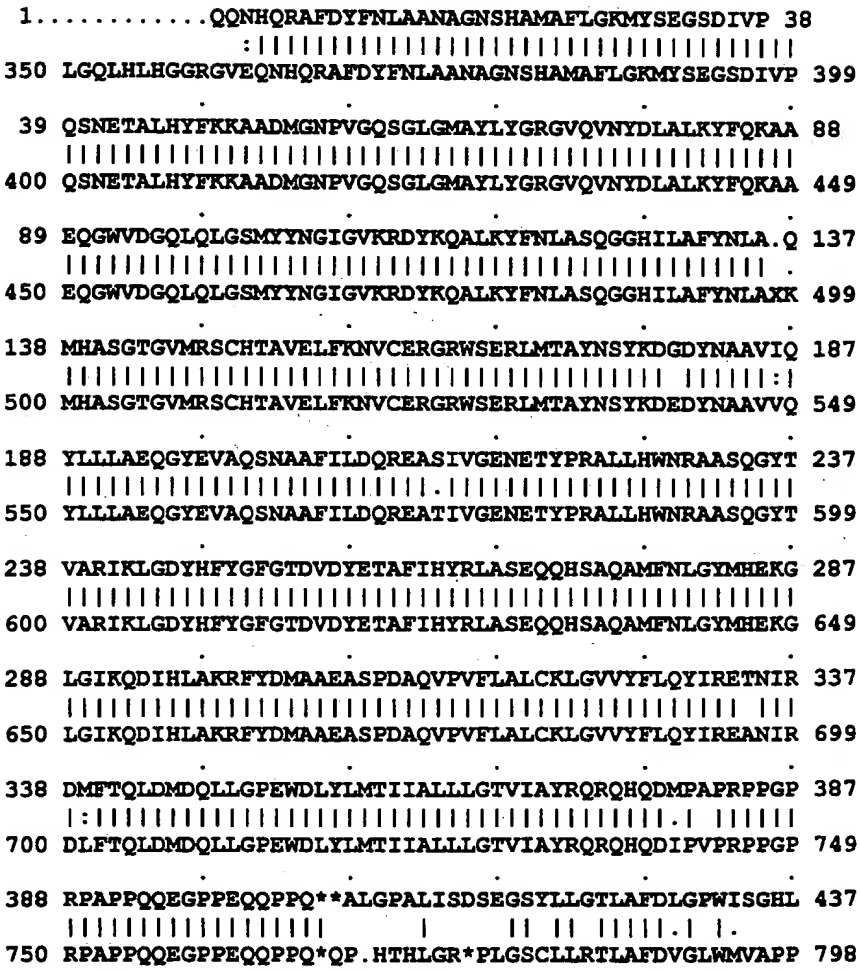
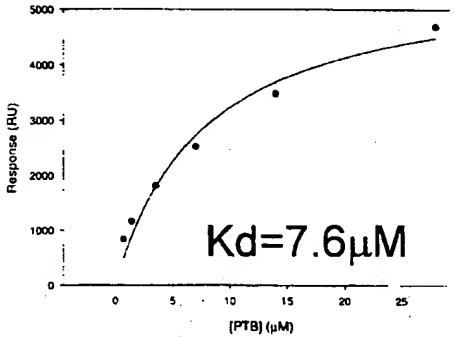

1.....QQNHQRAFDYFNLAANAGNSHAMAFLGKMYSEGSDIVP 38
350 LGQLHLHGGRGVEQNHQRAFDYFNLAANAGNSHAMAFLGKMYSEGSDIVP 399
39 QSNETALHYFKKAADMGNPVGQSGLGMAYLYGRGVQVNYDLALKYFQKAA 88
400 QSNETALHYFKKAADMGNPVGQSGLGMAYLYGRGVQVNYDLALKYFQKAA 449
89 EQGWVDGQLQLGSMYYNGIGVKRDYKQALKYFNLASQGGHILAFYNLA.Q 137
450 EQGWVDGQLQLGSMYYNGIGVKRDYKQALKYFNLASQGGHILAFYNLAXK 499
138 MHASGTGVMRSCHTAVELFKNVCERGRWSERLMTAYNSYKDGDYNAAVIQ 187
500 MHASGTGVMRSCHTAVELFKNVCERGRWSERLMTAYNSYKDEDYNAAVVQ 549
188 YLLLAEQGYEVAQSNAAFILDQREASIVGENETYPRALLHWNRAASQGYT 237
550 YLLLAEQGYEVAQSNAAFILDQREATIVGENETYPRALLHWNRAASQGYT 599
238 VARIKLGDYHFYGFGTDVDYETAFIHYRLASEQQHSAQAMFNLGYMHEKG 287
600 VARIKLGDYHFYGFGTDVDYETAFIHYRLASEQQHSAQAMFNLGYMHEKG 649
288 LGIKQDIHLAKRFYDMAAEASPDAQVPVFLALCKLGVVYFLQYIRETNIR 337
650 LGIKQDIHLAKRFYDMAAEASPDAQVPVFLALCKLGVVYFLQYIREANIR 699
338 DMFTQLDMDQLLGPEWDLYLMTIIALLLGTVIAYRQRQHQDMPAPRPPGP 387
700 DLFTQLDMDQLLGPEWDLYLMTIIALLLGTVIAYRQRQHQDIPVPRPPGP 749
388 RPAPPQQEGPPEQQPPQ**ALGPALISDSEGSYLLGTLAFDLGPWISGHL 437
750 RPAPPQQEGPPEQQPPQ*QP.HTHLGR*PLGSCLLRTLAFDVGLWMVAPP 798

Figure 8

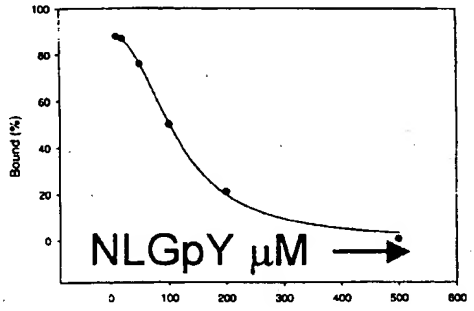
dNUMB-PTB domain binding to SEL-1L "NLGpY"-containing peptide

binding



Bia-core

Competition



Fluorescence Polarization

SEQ. ID. NO. 1

Partial Nucleotide Sequence of SEL-1L

AATCTTAA GTCTATTATC TCGTGTGATT TTTTTTTTTC TTACAGATGG

CATTGGAGTC AAGAGAGATT ATAAACAGGC CTTGAAGTAT TTTAATTTAG

CTTCTCAGGG AGGCCATATC TTGGCTTTCT ATAACCTAGC TCAGATGCAT

GCCAGTGGCA CCGGCGTGAT GCGATCATGT CACACTGCAG TGGAGGTAAG

GTCTTTCCTA CCAGCTTGTG

SEQ. ID. NO. 2

Partial Amino Acid Sequence of SEL-1L

ILSLLSR VIFFFLTDGI GVKRDYKQAL KYFNLASQGG

HILAFYNLAQ MHASGRGVMR SCHTAVEVR SFLPACV

SEQ. ID. NO. 3

Partial Nucleotide Sequence of SEL-1L

TTCCGGGAGAGAAGCATATCGGTACCTTCAGAAGGCAGCAGGCATGAATCACACCAAAGCC

CTGGAGAGAGTGTCCTATGCTCTCTTGTTTGGTGATTACCTCACACAGAATATCCAGGCAGCC

AAAGAGATGTTTGAGAAACTGACTGAGGAAGGGTCTCCCAAAGGACAGACTGGTCTTGGCTT

TCTCTACGCTTCTGGGCTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGTATATTATACTTTC

GGAGCTCTTGGAGGCAACCTGATAGCCCATATGATTTTGGGTTACCGCTACTGGGCTGGCATC

GGAGTCCTCCAGAGTTGTGAGTCGGCACTGACCCATTATCGTCTTGTTGCCAATCATGTTGCT

AGTGATATCTCCCTAACTGGAGGCTCTGTAGTCCAGAGAATACGGCTGCCCGATGAAGTGGA

AAACCCGGGGATGAACAGTGGGATGCTGGAAGAAGACCTGATTCAGTATTACCAGTTCCTAG

CTGAGAAGGGTGACGTCCAAGCACAGGTTGGCCTGGGACAGCTGCATCTGCATGGAGGGCGT

GGAGTAGAACAGAATCACCAGAGAGCGTTTGACTACTTCAACTTAGCAGCAAATGCTGGCAA

TTCACATGCTATGGCCTTCCTGGGAAAGATGTATTCTGAAGGAAGTGACATCGTACCTCAGAG

TAATGAGACGGCACTTCACTACTTTAAGAAAGCTGCTGACATGGGCAACCCCGTGGGACAGA

GCGGGCTTGGAATGGCCTACCTCTACGGAAGAGGCGTTCAAGTTAATTATGACCTGGCCCTCA

AG

NNNNNNNNNNNNNNNgap of about 36 nucleotidesNNNNNNNNNNNN

SEQ. ID. NO. 3 cont'd

Partial Nucleotide Sequence of SEL-1L

ATCTTAAGTCTATTATCTCGTGTGATTTTTTTTTTTCTTACAGATGGCATTGGAGTCAAGAGAG
ATTATAAACAGGCCTTGAAGTATTTTAATTTAGCTTCTCAGGGAGGCCATATCTTGGCTTTCTA
TAACCTAGCTCAGATGCATGCCAG
TGGNCGCGGCGTGATGCGATCATGTCACACTGCAGTGGAGGTAAGGTCTTTCCTACCAGCTTG

TGTG

NNNNNNNNNNNNNNgap of about 180 nucleotidesNNNNNNNNNN

GAAACTTACCCCAGAGCTTTACTGCATTGGAACAGGCCGCCTCCCAAGGTTACACTGTGGCTA
GAATTAAGCTTGGAGACTACCACTTCTATGGCTTTGGCACTGATGTGGATTATGAGACCGCAT
TTATTCATTACCGCCTGGCTTCTGAGCAGCAGCACAGCGCCCAAGCTATGTTTAACCTGGGCT
ACATGCACGAGAAGGGCCTAGGCATTAAACAGGACATTCACCTTGCAAAACGCTTTTATGAC
ATGGCAGCCGAAGCTAGCCCAGATGCACAAGTACCTGTGTTCCTCGCACTCTGCAAATTAGGT
GTCGTCTATTTCTTACAGTACATACGGGAAGCAAACATTCGAGATCTATTCACACAACTGGAT
ATGGACCAGCTTTTGGGACCCGAGTGGGACCTTTACCTCATGACCATCATTGCACTGCTCTTG
GGTACAGTCATAGCTTACAGGCAGCGACAGCACCAGGACATACCAGTTCCCGGCCCCCAGGG
CCACGGCCGGCTCCTCCCCAGCAGGAAGGACCACCAGAGCAGCAGCCACCACAGTAGCAGCC
ACACACTCACCTTGGTCGGTGACCACTGGGGAGCTGCTTGCTGAGAACGCTTGCATTTGATGT
AGGGCTCTGGATGGTGGCACCGCCTGGAAGAGGCGTGAGGCGTGTTGAATCCCAGAAGCTGC
TTAGAACCTGCTGCCTTCCTTTTCAGGGATGCGCGGCTCTCGGCAGAGCTGCAGTGAATGTTT
GTTTCAGTGCCATACGGAATGACAACTCTCAGTGGCTTTCCTCTTTCCTTTTTCCCGGAA

SEQ. ID. NO. 4

Partial Amino Acid Sequence of SEL-1L

PPSRSTVSISLISNSGREAYRYLQKAAGMNHTKALERVSYALLFGDYLTQNIQAAKEMFEKLTEEG

SPKGQTGLGFLYASGLGVNSSQAKALVYYTFGALGGNLIAHMILGYRYWAGIGVLQSCESALTH

YRLVANHVASDISLTGGSVVQRIRLPDEVENPGMNSGMLEEDLIQYYQFLAEKGDVQAQVGLGQ

LHLHGGRGVEQNHQRAFDYFNLAANAGNSHAMAFLGKMYSEGSDIVPQSNETALHYFKKAAD

MGNPVGQSGLGMAYLYGRGVQVNYDLALK

XXXXXXgap of 12 amino acidsXXXXXX

ILSLLSRVIFFFLTDGIGVKRDYKQALKYFNLASQGGHILAFYNLAQMHASXRGVMRSCHTAVEV

RSFLPACV XXXXXXgap of 60 amino acidsXXXXXXXXX

AASQGYTVARIKLGDYHFYGFGTDVDYETAFIHYRLASEQQHSAQAMFNLGYMHEKGLGIKQDI

HLAKRFYDMAAEASPDAQVPVFLALCKLGVVYFLQYIREANIRDLFTQLDMDQLLGPEW

LYLMTIIALLLGTVIAYRQRQHQDIPVPGPQGHGRLLPSRKDHQSSSHHSSSHTLTLVGDHWGAAC

SEQ. ID. NO. 5

Complete Mouse SEL-1L Nucleotide Sequence

```
   1  CGGGCTGCAG GAATTCCGGG TTGGTCGCGG CGACGGCGAC GGCGACGGCG
  51  AGGGCTCTCG ACCTTCGAGA GCAGGATGCA GGTCCGCGTA AGGCTGTCGT
 101  TGCTGCTGCT CTGCGCGGTG CTCCTGGGCT CGGCAGCCGC GACCTCGGAT
 151  GACAAAACTA ACCAGGATGA CTCCTTAGAT TCCAAGAGTT CTTTGCCCAC
 201  AGATGAGTCA GTGAAGGACC ACACCACCAC GGGCAAAGTA GTTGCTGGCC
 251  AGATATTTGT TGATTCTGAA GAAGCAGAAG TGGAATCCCT TCTTCAGGAC
 301  GAGGAAGATA GCTCCAAGAC CCAGGAGGAA GAGATCAGCT TTTTAGAATC
 351  TCCGAATCCA AGCAGCAAGA CCTACGAAGA ACTAAAGAGA GTGCGGAAGC
 401  CAGTCTTGAC TGCCATTGAA GCTGAAGAAG ATGCTGCCAA AAGACGACAG
 451  ATGCAGGAAG CAGAGATGAT CTATCAGGCC GGGATGAAGA TACTGAATGG
 501  AAGCAATAGG AAGAGCCAAA AGAGAGAAGC ATATCGGTAC CTTCAGAAGG
 551  CAGCAGGCAT GAATCACACC AAAGCCCTGG AGAGAGTGTC CTATGCTCTC
 601  TTGTTTGGTG ATTACCTCAC ACAGAATATC CAGGCAGCCA AAGAGATGTT
 651  TGAGAAACTG ACTGAGGAAG GGTCTCCCAA AGGACAGACT GGTCTTGGCT
 701  TTCTCTACGC TTCTGGGCTT GGTGTTAATT CAAGTCAGGC AAAGGCTCTT
 751  GTATATTATA CTTTCGGAGC TCTTGGAGGC AACCTGATAG CCCATATGAT
 801  TTTGGGTTAC CGCTACTGGG CTGGCATCGG AGTCCTCCAG AGTTGTGAGT
 851  CGGCACTGAC CCATTATCGT CTTGTTGCCA ATCATGTTGC TAGTGATATC
 901  TCCCTAACTG GAGGCTCTGT AGTCCAGAGA ATACGGCTGC CCGATGAAGT
```

```
 951  GGAAAACCCG GGGATGAACA GTGGGATGCT GGAAGAAGAC CTGATTCAGT
1001  ATTACCAGTT CCTAGCTGAG AAGGGTGACG TCCAAGCACA GGTTGGCCTG
1051  GGACAGCTGC ATCTGCATGG AGGGCGTGGA GTAGAACAGA ATCACCAGAG
1101  AGCGTTTGAC TACTTCAACT TAGCAGCAAA TGCTGGCAAT TCACATGCTA
1151  TGGCCTTCCT GGGAAAGATG TATTCTGAAG GAAGTGACAT CGTACCTCAG
1201  AGTAATGAGA CGGCACTTCA CTACTTTAAG AAAGCTGCTG ACATGGGCAA
1251  CCCCGTGGGA CAGAGCGGGC TTGGAATGGC CTACCTCTAC GGAAGAGGCG
1301  TTCAAGTTAA TTATGACCTG GCCCTCAAGT ATTTCCAGAA AGCTGCTGAG
1351  CAAGGCTGGG TGGACGGGCA GCTGCAGCTC GGCTCTATGT ACTACAATGG
1401  CATTGGAGTC AAGAGAGATT ATAAGCAGGC CTTGAAGTAT TTTAATCTGG
1451  CTTCTCAAGG AGGCCATATC TTGGCTTTCT ATAACCTCGC ACnnAAGATG
1501  CACGCCAGCG GCACAGGGGT GATGCGGTCC TGTCACACTG CAGTGGAGTT
1551  GTTTAAGAAT GTGTGTGAGC GAGGTCGCTG GTCAGAGAGA CTGATGACTG
1601  CCTACAACAG CTATAAGGAT GAGGACTACA ATGCTGCAGT GGTCCAGTAC
1651  CTCCTGCTGG CTGAGCAGGG CTACGAGGTG GCGCAGAGCA ACGCAGCCTT
1701  CATCCTCGAC CAGAGAGAAG CAACCATTGT AGGTGAGAAT GAAACTTACC
1751  CCAGAGCTTT ACTGCATTGG AACAGGGCCG CCTCCCAAGG TTACACTGTG
1801  GCTAGAATTA AGCTTGGAGA CTACCACTTC TATGGCTTTG GCACTGATGT
1851  GGATTATGAG ACCGCATTTA TTCATTACCG CCTGGCTTCT GAGCAGCAGC
```

```
1901  ACAGCGCCCA AGCTATGTTT AACCTGGGCT ACATGCACGA GAAGGGCCTA
1951  GGCATTAAAC AGGACATTCA CCTTGCAAAA CGCTTTTATG ACATGGCAGC
2001  CGAAGCTAGC CCAGATGCAC AAGTACCTGT GTTCCTCGCA CTCTGCAAAT
2051  TAGGTGTCGT CTATTTCTTA CAGTACATAC GGGAAGCAAA CATTCGAGAT
2101  CTATTCACAC AACTGGATAT GGACCAGCTT TTGGGACCCG AGTGGGACCT
2151  TTACCTCATG ACCATCATTG CACTGCTCTT GGGTACAGTC ATAGCTTACA
2201  GGCAGCGACA GCACCAGGAC ATACCAGTTC CCAGGCCCCC AGGGCCCACG
2251  GCCGGCTCCT CCCCAGCAGG AAGGACCACC AGAGCAGCAG CCACCACAGT
2301  AGCAGCCACA CACTCACCTT GGTCGGTGAC CACTGGGGAG CTGCTTGCTG
2351  AGAACGCTTG CATTTGATGT AGGGCTCTGG ATGGTGGCAC CGCCTGGAAG
2401  AGGCGTGAGG CGTGTTGAAT CCCAGAAGCT GCTTAGAACC TGCTGCCTTC
2451  CTTTTCAGGG ATGCGCGGCT CTCGGCAGAG CTGCAGTGAA TGTTTGTTTC
2501  AGTGCCATAC GGAATGACAA CTCTCAGTGG CTTTCCTCTT TCCTTTTTCT
2551  TTCTTCGTTC TGCAAACATA CACAAGACAC TCGAGTCATG TCTACTGTAC
2601  CCTGTCTTTC CTGAGAGACC TTTCGTCATC CTGTTGATGT GCATAACTTC
2651  TTACCTGTCT TCTAAGGTCC TGAGCATTGA CAACCTGGGG AAACCAGTTA
2701  GTTGCAAGTT TAACCCTAAA AACATGCTTA AGTCTGAGAC TAGAACTTGC
2751  CATTAAGTCA AAGCAAACAA GCCAAGCAGG ACGTCTTCAG GGTGTTTGGG
2801  AAATGCAGTA ACTGGGAAAT TTTACCAGTC TACATGCCAG GTCATATGTT
```

2851 CGGCACTCTT ATTTTAAAAG GGATTATTGA TCGGATTCCT GTACTGTGAC

2901 TCTGACACTT CCTAGCAAAC TTCTCTTCCT CAAAGACTGA AAGGCTTTGC

2951 TAAAGACTTC TTCGTGATGT CACATCCCTG CCATGGAGTA GAAGTCAAAC

3001 TCAGTCAGCT GACTCCGTGC TCCCACAGAG GACTTTGCCC CAGCTACTCA

3051 TGTGAGTTGC ATATACAGAG CCCTAGTCCA GACGCCAGAG GATGATCTAA

3101 ATGGAGAAAT CAGAAAACAG CCTTAACCAA CTTGAGGTGG GACCTGGGAC

3151 ATGCCCTCCC TCGGGTAGAT TCGTTGCTTC CCGGATCTCA TTTTACTTCA

3201 CATTTCTTAT GGCACAGAGA ACAGACAACC CTCCCAAAGT CTTCAAATAT

3251 GAAATGCTAT TGTCAAGATT GGCAGGTGGT GACTTCCTTT GTAATACTAA

3301 AATACTAAAA TACTCAGACA ATGGAGGATG GGTTAGACCC AATGCATGGC

3351 AAAACCAAAG GCCTGGGTGC ACCGTTGTGT ATGCTGCGGG AGCTTTTTAT

3401 TTCTGACGGT TTGATGTGGC TGTAGAAAGT CAGCCACTGA GTAGCTGTTG

3451 TGTTATCGCC CTTCTGTTGT AAACATTAGG AAATTAAAGC AGTTACACCC

3501 AGCATTGCCT GCTTCAGTAA GAGGCGAAAA CAACGACTGA GAGCTTCACA

3551 TAATCACCCC AAAGCCCAAG gGAATTTGTG GTCTGTGTAA TCAGATCTGT

3601 GGATGGTAGC GTTGGGTAGC CAGCACCCCA TGGGAGTCAC GCCCGGAATT

3651 CCTGCAGCCC GGGGATC

SEQ. ID. NO. 6

Translation of Mouse SEL-1L Complete Sequence

```
  1  RAAGIPGWSR RRRRRRRGLS TFESRMQVRV RLSLLLLCAV LLGSAAATSD
 51  DKTNQDDSLD SKSSLPTDES VKDHTTTGKV VAGQIFVDSE EAEVESLLQD
101  EEDSSKTQEE EISFLESPNP SSKTYEELKR VRKPVLTAIE AEEDAAKRRQ
151  MQEAEMIYQA GMKILNGSNR KSQKREAYRY LQKAAGMNHT KALERVSYAL
201  LFGDYLTQNI QAAKEMFEKL TEEGSPKGQT GLGFLYASGL GVNSSQAKAL
251  VYYTFGALGG NLIAHMILGY RYWAGIGVLQ SCESALTHYR LVANHVASDI
301  SLTGGSVVQR IRLPDEVENP GMNSGMLEED LIQYYQFLAE KGDVQAQVGL
351  GQLHLHGGRG VEQNHQRAFD YFNLAANAGN SHAMAFLGKM YSEGSDIVPQ
401  SNETALHYFK KAADMGNPVG QSGLGMAYLY GRGVQVNYDL ALKYFQKAAE
451  QGWVDGQLQL GSMYYNGIGV KRDYKQALKY FNLASQGGHI LAFYNLAxKM
501  HASGTGVMRS CHTAVELFKN VCERGRWSER LMTAYNSYKD EDYNAAVVQY
551  LLLAEQGYEV AQSNAAFILD QREATIVGEN ETYPRALLHW NRAASQGYTV
601  ARIKLGDYHF YGFGTDVDYE TAFIHYRLAS EQQHSAQAMF NLGYMHEKGL
651  GIKQDIHLAK RFYDMAAEAS PDAQVPVFLA LCKLGVVYFL QYIREANIRD
701  LFTQLDMDQL LGPEWDLYLM TIIALLLGTV IAYRQRQHQD IPVPRPPGPT
751  AGSSPAGRTT RAAATTVAAT HSPWSVTTGE LLAENACI*
```

SEQ. ID. NO. 7

Complete Mouse SEL-1L Nucleotide Sequence Splice Variant

1 CGGGCTGCAG GAATTCCGGG TTGGTCGCGG CGACGGCGAC GGCGACGGCG

51 AGGGCTCTCG ACCTTCGAGA GCAGGATGCA GGTCCGCGTA AGGCTGTCGT

101 TGCTGCTGCT CTGCGCGGTG CTCCTGGGCT CGGCAGCCGC GACCTCGGAT

151 GACAAAACTA ACCAGGATGA CTCCTTAGAT TCCAAGAGTT CTTTGCCCAC

201 AGATGAGTCA GTGAAGGACC ACACCACCAC GGGCAAAGTA GTTGCTGGCC

251 AGATATTTGT TGATTCTGAA GAAGCAGAAG TGGAATCCCT TCTTCAGGAC

301 GAGGAAGATA GCTCCAAGAC CCAGGAGGAA GAGATCAGCT TTTTAGAATC

351 TCCGAATCCA AGCAGCAAGA CCTACGAAGA ACTAAAGAGA GTGCGGAAGC

401 CAGTCTTGAC TGCCATTGAA GCTGAAGAAG ATGCTGCCAA AAGACGACAG

451 ATGCAGGAAG CAGAGATGAT CTATCAGGCC GGGATGAAGA TACTGAATGG

501 AAGCAATAGG AAGAGCCAAA AGAGAGAAGC ATATCGGTAC CTTCAGAAGG

551 CAGCAGGCAT GAATCACACC AAAGCCCTGG AGAGAGTGTC CTATGCTCTC

601 TTGTTTGGTG ATTACCTCAC ACAGAATATC CAGGCAGCCA AAGAGATGTT

651 TGAGAAACTG ACTGAGGAAG GGTCTCCCAA AGGACAGACT GGTCTTGGCT

701 TTCTCTACGC TTCTGGGCTT GGTGTTAATT CAAGTCAGGC AAAGGCTCTT

751 GTATATTATA CTTTCGGAGC TCTTGGAGGC AACCTGATAG CCCATATGAT

801 TTTGGGTTAC CGCTACTGGG CTGGCATCGG AGTCCTCCAG AGTTGTGAGT

851 CGGCACTGAC CCATTATCGT CTTGTTGCCA ATCATGTTGC TAGTGATATC

901 TCCCTAACTG GAGGCTCTGT AGTCCAGAGA ATACGGCTGC CCGATGAAGT

951 GGAAAACCCG GGGATGAACA GTGGGATGCT GGAAGAAGAC CTGATTCAGT

1001 ATTACCAGTT CCTAGCTGAG AAGGGTGACG TCCAAGCACA GGTTGGCCTG

1051 GGACAGCTGC ATCTGCATGG AGGGCGTGGA GTAGAACAGA ATCACCAGAG

1101 AGCGTTTGAC TACTTCAACT TAGCAGCAAA TGCTGGCAAT TCACATGCTA

1151 TGGCCTTCCT GGGAAAGATG TATTCTGAAG GAAGTGACAT CGTACCTCAG

1201 AGTAATGAGA CGGCACTTCA CTACTTTAAG AAAGCTGCTG ACATGGGCAA

1251 CCCCGTGGGA CAGAGCGGGC TTGGAATGGC CTACCTCTAC GGAAGAGGCG

1301 TTCAAGTTAA TTATGACCTG GCCCTCAAGT ATTTCCAGAA AGCTGCTGAG

1351 CAAGGCTGGG TGGACGGGCA GCTGCAGCTC GGCTCTATGT ACTACAATGG

1401 CATTGGAGTC AAGAGAGATT ATAAGCAGGC CTTGAAGTAT TTTAATCTGG

SEQ. ID. NO. 7 cont'd

Complete Mouse SEL-1L Nucleotide Sequence Splice Variant

1451 CTTCTCAAGG AGGCCATATC TTGGCTTTCT ATAACCTCGC ACnnAAGATG

1501 CACGCCAGCG GCACAGGGGT GATGCGGTCC TGTCACACTG CAGTGGAGTT

1551 GTTTAAGAAT GTGTGTGAGC GAGGTCGCTG GTCAGAGAGA CTGATGACTG

1601 CCTACAACAG CTATAAGGAT GAGGACTACA ATGCTGCAGT GGTCCAGTAC

1651 CTCCTGCTGG CTGAGCAGGG CTACGAGGTG GCGCAGAGCA ACGCAGCCTT

1701 CATCCTCGAC CAGAGAGAAG CAACCATTGT AGGTGAGAAT GAAACTTACC

1751 CCAGAGCTTT ACTGCATTGG AACAGGGCCG CCTCCCAAGG TTACACTGTG

1801 GCTAGAATTA AGCTTGGAGA CTACCACTTC TATGGCTTTG GCACTGATGT

1851 GGATTATGAG ACCGCATTTA TTCATTACCG CCTGGCTTCT GAGCAGCAGC

1901 ACAGCGCCCA AGCTATGTTT AACCTGGGCT ACATGCACGA GAAGGGCCTA

1951 GGCATTAAAC AGGACATTCA CCTTGCAAAA CGCTTTTATG ACATGGCAGC

2001 CGAAGCTAGC CCAGATGCAC AAGTACCTGT GTTCCTCGCA CTCTGCAAAT

2051 TAGGTGTCGT CTATTTCTTA CAGTACATAC GGGAAGCAAA CATTCGAGAT

2101 CTATTCACAC AACTGGATAT GGACCAGCTT TTGGGACCCG AGTGGGACCT

2151 TTACCTCATG ACCATCATTG CACTGCTCTT GGGTACAGTC ATAGCTTACA

2201 GGCAGCGACA GCACCAGGAC ATACCAGTTC CCAGGCCCCC AGGGCCACGG

2251 CCGGCTCCTC CCCAGCAGGA AGGACCACCA GAGCAGCAGC CACCACAGTA

2301 GCAGCCACAC ACTCACCTTG GTCGGTGACC ACTGGGGAGC TGCTTGCTGA

2351 GAACGCTTGC ATTTGATGTA GGGCTCTGGA TGGTGGCACC GCCTGGAAGA

2401 GGCGTGAGGC GTGTTGAATC CCAGAAGCTG CTTAGAACCT GCTGCCTTCC

2451 TTTTCAGGGA TGCGCGGCTC TCGGCAGAGC TGCAGTGAAT GTTTGTTTCA

2501 GTGCCATACG GAATGACAAC TCTCAGTGGC TTTCCTCTTT CCTTTTTCTT

2551 TCTTCGTTCT GCAAACATAC ACAAGACACT CGAGTCATGT CTACTGTACC

2601 CTGTCTTTCC TGAGAGACCT TTCGTCATCC TGTTGATGTG CATAACTTCT

2651 TACCTGTCTT CTAAGGTCCT GAGCATTGAC AACCTGGGGA AACCAGTTAG

2701 TTGCAAGTTT AACCCTAAAA ACATGCTTAA GTCTGAGACT AGAACTTGCC

2751 ATTAAGTCAA AGCAAACAAG CCAAGCAGGA CGTCTTCAGG GTGTTTGGGA

2801 AATGCAGTAA CTGGGAAATT TTACCAGTCT ACATGCCAGG TCATATGTTC

SEQ. ID. NO. 7 cont'd

Complete Mouse SEL-1L Nucleotide Sequence Splice Variant

2851 GGCACTCTTA TTTTAAAAGG GATTATTGAT CGGATTCCTG TACTGTGACT

2901 CTGACACTTC CTAGCAAACT TCTCTTCCTC AAAGACTGAA AGGCTTTGCT

2951 AAAGACTTCT TCGTGATGTC ACATCCCTGC CATGGAGTAG AAGTCAAACT

3001 CAGTCAGCTG ACTCCGTGCT CCCACAGAGG ACTTTGCCCC AGCTACTCAT

3051 GTGAGTTGCA TATACAGAGC CCTAGTCCAG ACGCCAGAGG ATGATCTAAA

3101 TGGAGAAATC AGAAAACAGC CTTAACCAAC TTGAGGTGGG ACCTGGGACA

3151 TGCCCTCCCT CGGGTAGATT CGTTGCTTCC CGGATCTCAT TTTACTTCAC

3201 ATTTCTTATG GCACAGAGAA CAGACAACCC TCCCAAAGTC TTCAAATATG

3251 AAATGCTATT GTCAAGATTG GCAGGTGGTG ACTTCCTTTG TAATACTAAA

3301 ATACTAAAAT ACTCAGACAA TGGAGGATGG GTTAGACCCA ATGCATGGCA

3351 AAACCAAAGG CCTGGGTGCA CCGTTGTGTA TGCTGCGGGA GCTTTTTATT

3401 TCTGACGGTT TGATGTGGCT GTAGAAAGTC AGCCACTGAG TAGCTGTTGT

3451 GTTATCGCCC TTCTGTTGTA AACATTAGGA AATTAAAGCA GTTACACCCA

3501 GCATTGCCTG CTTCAGTAAG AGGCGAAAAC AACGACTGAG AGCTTCACAT

3551 AATCACCCCA AAGCCCAAGg GAATTTGTGG TCTGTGTAAT CAGATCTGTG

3601 GATGGTAGCG TTGGGTAGCC AGCACCCCAT GGGAGTCACG CCCGGAATTC

3651 CTGCAGCCCG GGGATC

SEQ. ID. NO. 8

Translation of Mouse SEL-1L Complete Sequence in ID. NO. 7

RAAGIPGWSR RRRRRRRGLS TFESRMQVRV RLSLLLLCAV LLGSAAATSD
DKTNQDDSLD SKSSLPTDES VKDHTTTGKV VAGQIFVDSE EAEVESLLQD
EEDSSKTQEE EISFLESPNP SSKTYEELKR VRKPVLTAIE AEEDAAKRRQ
MQEAEMIYQA GMKILNGSNR KSQKREAYRY LQKAAGMNHT KALERVSYAL
LFGDYLTQNI QAAKEMFEKL TEEGSPKGQT GLGFLYASGL GVNSSQAKAL
VYYTFGALGG NLIAHMILGY RYWAGIGVLQ SCESALTHYR LVANHVASDI
SLTGGSVVQR IRLPDEVENP GMNSGMLEED LIQYYQFLAE KGDVQAQVGL
GQLHLHGGRG VEQNHQRAFD YFNLAANAGN SHAMAFLGKM YSEGSDIVPQ
SNETALHYFK KAADMGNPVG QSGLGMAYLY GRGVQVNYDL ALKYFQKAAE
QGWVDGQLQL GSMYYNGIGV KRDYKQALKY FNLASQGGHI LAFYNLAXKM
HASGTGVMRS CHTAVELFKN VCERGRWSER LMTAYNSYKD EDYNAAVVQY
LLLAEQGYEV AQSNAAFILD QREATIVGEN ETYPRALLHW NRAASQGYTV
ARIKLGDYHF YGFGTDVDYE TAFIHYRLAS EQQHSAQAMF NLGYMHEKGL
GIKQDIHLAK RFYDMAAEAS PDAQVPVFLA LCKLGVVYFL QYIREANIRD
LFTQLDMDQL LGPEWDLYLM TIIALLLGTV IAYRQRQHQD IPVPRPPGPR
PAPPQQEGPP EQQPPQ*

SEQ. ID. NO. 9

Human SEL-1L Partial Nucleotide Sequence

Sel-1l human cDNA 19a clone 1, consensus.
Primer T7.

1 ATGGATATCT GCAGAATTCG GCTTGGTAAT GGTAGCGACC GGCGCTCAGC

51 TGGAATTCGC GGCCGCGTCG ACAACAGAAT CATCAGAGAG CATTTGACTA

101 CTTCAATTTA GCAGCAAATG CTGGCAATTC ACATGCCATG GCCTTTTTGG

151 GAAAGATGTA TTCGGAAGGA AGTGACATTG TACCTCAGAG TAATGAGACA

201 GCTCTCCACT ACTTTAAGAA AGCTGCTGAC ATGGGCAACC CAGTTGGACA

251 GAGTGGGCTT GGAATGGCCT ACCTCTATGG GAGAGGAGTT CAAGTTAATT

301 ATGATCTAGC CCTTAAGTAT TTCCAGAAAG CTGCTGAACA AGGCTGGGTG

351 GATGGGCAGC TACAGCTTGG TTCCATGTAC TATAATGGCA TTGGAGTCAA

401 GAGAGATTAT AAACAGGCCT TGAAGTATTT TAATTTAGCT TCTCAGGGAG

451 GCCATATCTT GGCTTTCTAT AACCTAGCTC AGATGCATGC CAGTGGCACC

501 GGCGTGATGC GATCATGTCA CACTGCAGTG GAGTTGTTTA AGAATGTATG

551 TGAACGAGGC CGTTGGTCTG AAAGGCTTAT GACTGCCTAT AACAGCTATA

601 AAGATGGCGA TTACAATGCT GCAGTGATCC AGTACCTCCT CCTGGCTGAA

651 CAGGGCTATG AAGTGGCACA AAGCAATGCA GCCTTTATTC TTGATCAGAG

701 AGAAGCAAGC ATTGTAGGTG AGAATGAAAC TTATCCCAGA GCTTTGCTAC

751 ATTGGAACAG GGCCGCCTCT CAAGGCTATA CTGTGGCTAG AATTAAGCTC

801 GGAGACTACC ATTTCTATGG GTTTGGCACC GATGTAGATT ATGAAACTGC

851 ATTTATTCAT TACCGTCTGG CTTCTGAGCA GCAACACAGT GCACAAGCTA

901 TGTTTAATCT GGGATATATG CATGAGAAAG GACTGGGCAT TAAACAGGAT

951 ATTCACCTTG CGAAACGTTT TTATGACATG GCAGCTGAAG CCAGCCCAGA

1001 TGCACAAGTT CCAGTCTTCC TAGCCCTCTG CAAATTGGGC GTCGTCTATT

1051 TCTTGCAGTA CATACGGGAA ACAAACATTC GAGATATGTT CACCCAACTT

1101 GATATGGACC AGCTTTTGGG ACCTGAGTGG GACCTTTACC TCATGACCAT

1151 CATTGCGCTG CTGTTGGGAA CAGTCATAGC TTACAGGCAA AGGCAGCACC

1201 AAGACATGCC TGCACCCAGG CCTCCAGGGC CACGGCCAGC TCCACCCCAG

1251 CAGGAGGGGC CACCAGAGCA GCAGCCACCA CAGTAATAGG CACTGGGTCC

1301 AGCCTTGATC AGTGACAGCG AAGGAAGTTA TCTGCTGGGA ACACTTGCAT

1351 TTGATTTAGG ACCTTGGATC AGTGGTCACC TCCCAGAAGA GGCACGGCAC

1401 AAGGAAGCAT TGAATTCCTA AAGCTGCTTA GAATCTGATG CCTTTATTTT

1451 CAGGGATAAG TAACTCTTAC CTAAACTGAG CTGAATGTTT GTTTCAGTGC

1501 CATATGGAAT AACAACTTTC AGTGGCTTTT TTTTTCTTTT CTGGAAACAT

1551 ATGCGAGACA CTCAGAGTAA TGTCTACTGT ATCCAGCTAT CTTTCTTGGA

1601 TCCTTTTGGT CATTATTTCA GTGTGCATAA GTTCTTAATG TCAACCATCT

1651 TTAAGGTATT GTGCATCGAC ACTAAAAACT GATCAGTGTA AAAAGGAAAA

1701 CCCAGTTGCA AGTTTAAACG TGTTCGAAAG TCTGAAAATA GAACTTGCCT

1751 TTTAAGTTAA AAAAAAAAAA AGCTATCTTG AAAATGTTTT GGAACTGCGA

1801 TAACTGAGAA ACTCTTACCA GTCCACATGC AATTAGACAT ATTCAGCATA

1851 TTTGTTATTT TAAAAGGGAG GGTTGGGAGG TTTCTTATTG GTGATTGTCA

1901 CACGGTATAC CATACTCCTC TCCTTCAAAG AATGAAAGGC CTTGTTAAGG

1951 AGTTTTTTGT GAGCTTTACT TCTTTGGAAT GGAATATACT TATGCAAAAC

2001 CTTGTGAACT GACTCCTTGC ACTAACGCGA GTTTGCCCCA CCTACTCTGT

2051 AATTTGCTTG TTTGTTTTGA ATATACAGAG CCTTGATCCA GAAGCCAGAG

2101 GATGGACTA

SEQ. ID. NO. 10

Translation of Human SEL-1L Partial Sequence

Sel-1l human cDNA 19a clone 1, consensus.
Primer T7.

Pep Length: 185 1

1 QQNHQRAFDY FNLAANAGNS HAMAFLGKMY SEGSDIVPQS NETALHYFKK

51 AADMGNPVGQ SGLGMAYLYG RGVQVNYDLA LKYFQKAAEQ GWVDGQLQLG

101 SMYYNGIGVK RDYKQALKYF NLASQGGHIL AFYNLAQMHA SGTGVMRSCH

151 TAVELFKNVC ERGRWSERLM TAYNSYKDGD YNAAVIQYLL LAEQGYEVAQ

201 SNAAFILDQR EASIVGENET YPRALLHWNR AASQGYTVAR IKLGDYHFYG

251 FGTDVDYETA FIHYRLASEQ QHSAQAMFNL GYMHEKGLGI KQDIHLAKRF

301 YDMAAEASPD AQVPVFLALC KLGVVYFLQY IRETNIRDMF TQLDMDQLLG

351 PEWDLYLMTI IALLLGTVIA YRQRQHQDMP APRPPGPRPA PPQQEGPPEQ

401 QPPQ**